



МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

Национальный исследовательский
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

**ДЕСЯТАЯ
СИБИРСКАЯ КОНФЕРЕНЦИЯ
ПО ПАРАЛЛЕЛЬНЫМ И
ВЫСОКОПРОИЗВОДИТЕЛЬНЫМ
ВЫЧИСЛЕНИЯМ**

СБОРНИК СТАТЕЙ

Томск, 5–7 октября 2021 г.



ТОМСК
«Издательство НТЛ»
2021

УДК 519.6
ББК 22.19
ДЗ77

ДЗ77 **Десятая Сибирская конференция по параллельным и высокопроизводительным вычислениям: сборник статей** (Томск, 5–7 октября 2021 г.) / под ред. А.В. Старченко. – Томск: Изд-во НТЛ, 2021. – 122 с.

ISBN 978-5-89503-652-5

Включены материалы Десятой Сибирской конференции по параллельным и высокопроизводительным вычислениям, проходившей 5–7 октября 2021 г. в Томском государственном университете. Конференция посвящена 100-летию со дня рождения выпускника физико-математического факультета ТГУ, Героя Социалистического Труда академика Н.Н. Яненко.

Рассмотрены актуальные проблемы организации параллельных вычислений на многопроцессорных системах, современное состояние и перспективы развития методов параллельных вычислений. Для студентов, аспирантов, преподавателей, научных работников, желающих изучить и практически использовать в научной работе высокопроизводительные вычислительные ресурсы.

УДК 519.6
ББК 22.19

Организаторы:

Национальный исследовательский Томский государственный университет; механико-математический факультет НИ ТГУ; Региональный научно-образовательный математический центр (НОМЦ) НИ ТГУ.

Редакционная коллегия:

Старченко А.В., докт. физ.-мат. наук, зав. кафедрой ВМиКМ ММФ ТГУ; *Берцун В.Н.*, канд. физ.-мат. наук, доцент ММФ ТГУ; *Данилкин Е.А.*, канд. физ.-мат. наук, доцент ММФ ТГУ.

Конференция и издание сборника поддержаны Региональным научно-образовательным математическим центром Томского государственного университета за счет средств субсидии в целях оказания федеральным государственным учреждениям дополнительной государственной поддержки (соглашение № 075-02-2021-1392).

ISBN 978-5-89503-652-5

© ООО «Издательство НТЛ», 2021

СОДЕРЖАНИЕ

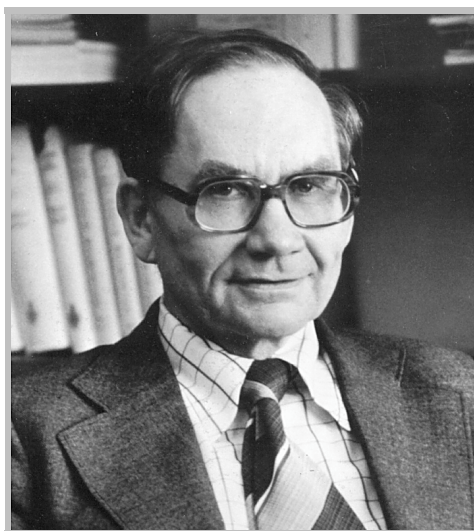
Берцун В.Н. Страницы жизни академика Н.Н. Яненко	4
Старченко А.В., Данилкин Е.А., Проханов С.А., Лещинский Д.В. К выбору гибридной схемы распараллеливания для численного решения уравнений пространственной мезомасштабной метеорологической модели атмосферного пограничного слоя	7
Данилкин Е.А., Старченко А.В., Каратаева Е.А., Юмин К.В., Дель И.В., Смиян Н.С., Лещинский Д.В. Параллельная реализация некоторых итерационных методов для решения системы линейных алгебраических уравнений в мезомасштабной метеорологической модели атмосферного пограничного слоя	18
Перепёлкин В.А., Иванов М.И. Повышение производительности LuNA-программ на основе воспроизведения трасс	29
Ботыгин И.А., Шерстнёв В.С., Шерстнёва А.И. Исследование возможностей языка программирования R для обработки больших данных	37
Leshchinskiy D.V., Danilkin E.A., Starchenko A.V. Influence of the pollution source location on the concentration of the impurity in street canyon	47
Зинченко В.И., Гольдин В.Д., Лаева В.И. Расчет сопряженного нестационарного теплообмена затупленных тел при их сверхзвуковом обтекании	53
Гурина Е.И. Математические модели и численные методы для расчета взаимодействия двух осевых вентиляторов	70
Хамидов А.Н., Михайлов М.Д. Исследование математической модели типа «хищник – жертва» с учетом поискового поведения хищника	80
Дель И.В., Старченко А.В. Прогноз приземной температуры воздуха на основе модели рекуррентной нейронной сети типа LSTM	91
Шумилов Б.М. О построении предобуславливателей для четно-нечетного расщепления сплайн-вейвлетов	101
Берцун В.Н. О кратномасштабном вейвлет-анализе сеточных функций	112

Страницы жизни академика Н.Н. Яненко

В.Н. Берцун

Томский государственный университет, г. Томск, Россия

Николай Николаевич Яненко родился 22 мая 1921 г. в г. Куйбышеве (г. Каинск) Новосибирской области [1]. В 1923 г. Николай с родителями, двумя старшими братьями и двумя сестрами переехали в г. Новосибирск. С восьмого класса он учился в железнодорожной школе № 1 г. Новосибирска, которую окончил с отличием в 1939 г. Уже в юные годы Николай отличался от своих сверстников феноменальной памятью. После окончания школы он поступает на ФМФ ТГУ.



ЯНЕНКО
Николай Николаевич
(1921–1984)

Как отличник, Николай Яненко был зачислен без экзаменов на специальность «Математика». Университет он окончил досрочно с отличием в 1942 г. С декабря 1942 г. он был призван в действующую армию, из которой был демобилизован в звании лейтенанта в декабре 1945 г. и, приехав в Москву, стал готовиться к поступлению в аспирантуру МГУ.

В 1946 г. Николай Николаевич поступает в аспирантуру ММФ МГУ к профессору Петру Константиновичу Рашевскому. Через три года он успешно защитил кандидатскую, а в 1954 г. – докторскую диссертации по дифференциальной топологии.

В 1948 г. произошло важное событие в научной биографии Н.Н. Яненко: он начал работать в группе академика Андрея Николаевича Тихонова, которому была поручена организация вычислительной лаборатории для проведения расчетов процесса взрыва атомной (а потом и водородной) бомбы. За активное участие в этих проектах, по представлению А.Н. Тихонова, Н.Н. Яненко был награжден денежной премией (1949 г.), а позднее был удостоен Сталинской премии третьей степени (1953 г.).

В 1955 г. на Урале был открыт новый ядерный центр (ВНИИТФ, г. Снежинск), где Н.Н. Яненко в 34 года был назначен начальником математического отделения. Предвидя стремительное развитие вычислительной техники и вычислительной математики, Н.Н. Яненко активизировал создание экономичных разностных схем для расчета многомерных задач механики сплошной среды.

В 1957 г. в Новосибирске было создано СО АН СССР. По приглашению директора ВЦ СО РАН Гурия Ивановича Марчука в октябре 1963 г. Н.Н. Яненко вместе с семьей переехал в Новосибирск на работу в Вычислительном центре СО АН СССР. Он назначается заведующим отделом численных методов механики сплошной среды этого центра. В 1966 г. Николай Николаевич был избран в члены-корреспонденты АН СССР, а в 1970 г. становится академиком АН СССР по отделению механики и процессов управления. С 1976 г. в течение 8 лет он был директором Института теоретической и прикладной механики СО АН СССР.

Одним из главных достижений Н.Н. Яненко в вычислительной математике является создание (1957–1963 гг.) «метода дробных шагов», позволяющего свести решение многомерной задачи к последовательному решению соответствующей цепочки одномерных задач. Результаты этой работы были сконцентрированы в монографии Н.Н. Яненко «Метод дробных шагов решения многомерных задач математической физики», которая вышла в 1967 г. и была переведена на немецкий, французский и английский языки.

Николай Николаевич способствовал созданию стройной системы подготовки кадров для ВЦ СО АН СССР и ММФ ТГУ, в соответствии с которой первые три курса студенты механико-математического факультета учились в ТГУ, а затем проходили специализацию и заканчивали

обучение в НГУ. Благодаря его поддержке в лабораториях НИИ ПММ при ТГУ использовались новые постановки задач и передовые математические технологии.

Николай Николаевич много раз приезжал в ТГУ для чтения лекций. Он поддержал создание на ММФ ТГУ кафедры физической механики (1977 г.) и отдела механики реагирующих сред в НИИ ПММ при ТГУ, проводил консультации, устраивал Всесоюзные школы-семинары по численным методам механики вязкой жидкости и механике реагирующих сред. Николай Николаевич был одним из инициаторов создания пакетов прикладных программ, параллельных вычислительных технологий, метода дифференциального приближения, исследований по интервальному анализу и др.

В день шестидесятилетия 22 мая 1981 г. ему было присвоено звание Героя Социалистического Труда. Николай Николаевич имеет более 350 научных работ и 15 монографий, в которых он является автором либо соавтором.

Вся его жизнь, сибирский характер, талант организатора, беспредельная преданность науке и своему Отечеству, неутомимый труд и высокая принципиальность являются яркой страницей в истории мировой и отечественной математики, вычислительной механики и суперкомпьютерных технологий, в истории первого на просторах Сибири Томского государственного университета.

СПИСОК ЛИТЕРАТУРЫ

1. Берцун В.Н., Косова Е.С. Страницы жизни выпускника трех университетов (К 100-летию со дня рождения академика Н.Н. Яненко) // Вестник Томского государственного университета. Математика и механика. – 2021. – № 73. С. 135–145. – DOI: 10.17223/19988621/73/12.

Берцун Владимир Николаевич, канд. физ.-мат. наук, доцент кафедры вычислительной математики и компьютерного моделирования ММФ НИ ТГУ; bvn@math.tsu.ru.

К выбору гибридной схемы распараллеливания для численного решения уравнений пространственной мезомасштабной метеорологической модели атмосферного пограничного слоя*

А.В. Старченко¹, Е.А. Данилкин^{1,2},
С.А. Проханов¹, Д.В. Лещинский^{1,2}

¹Томский государственный университет, г. Томск, Россия

²Институт оптики атмосферы им. В.Е. Зуева СО РАН, г. Томск, Россия

Рассматривается параллельный алгоритм численного решения обобщенного трехмерного дифференциального уравнения конвективно-диффузионного переноса для мезомасштабной метеорологической модели. Проведено сравнение эффективности применения следующих технологий параллельного программирования: Message Passing Interface (MPI), Open Multi-Processing (OpenMP) и их комбинации MPI+OpenMP. Показано, что при использовании до 145 вычислительных ядер наиболее перспективной технологией распараллеливания является MPI.

Ключевые слова: MPI, OpenMP, параллельные вычисления, метеорологические модели.

Одной из актуальных проблем как фундаментальных, так и прикладных наук является создание математического и суперкомпьютерного программного обеспечения для информационных систем мониторинга и прогнозирования состояния приземного слоя атмосферы над населенными пунктами и крупными транспортными узлами. В ТГУ разрабатывается и используется для прогноза локальной погоды мезомасштабная метеорологическая модель высокого разрешения TSUNM3 [1].

Современные математические модели представляют собой систему из нескольких конвективно-диффузионных уравнений, дополненную замыкающими алгебраическими соотношениями, а также начальными и граничными условиями. Для решения таких систем применяются методы конечных разностей, конечного объема или конечных элементов. Приближенное решение ищется на подробной сетке, состоящей из не-

* Исследование выполнено за счет гранта Российского научного фонда (проект №19-71-20042).

скольких миллионов вычислительных узлов или ячеек. Причем при оперативном моделировании атмосферных процессов необходимо в кратчайшие сроки представить результаты прогностического моделирования. Хорошую возможность ускорить процесс получения численного решения дает применение многопроцессорной и многоядерной вычислительной техники [2, 3].

Данная работа посвящена построению гибридной схемы распараллеливания алгоритма для численного решения уравнений пространственной мезомасштабной метеорологической модели атмосферного пограничного слоя. Гибридный алгоритм построен как комбинация двух технологий параллельного программирования MPI и OpenMP. Библиотека MPI используется для взаимодействия между процессами, а распараллеливание в рамках каждого процесса выполняется с использованием технологии OpenMP.

Математическая постановка задачи и численный метод ее решения

Мезомасштабная модель TSUNM3 (Tomsk State University Nonhydrostatic Mesoscale Meteorology Model) прогнозирует составляющие скорости ветра и температурно-влажностные характеристики в пограничном слое атмосферы на 50 вертикальных уровнях (до 10 км) для площади 200×200 км с вложенной областью 50×50 км (шаг сетки 1 км с центром в г. Томск) [1]. Инициализация модели проводится на основе результатов численного прогноза по оперативной глобальной модели ПЛАВ Гидрометцентра РФ [4].

Математическая формулировка модели TSUNM3 включает одиннадцать нестационарных трехмерных неоднородных уравнений «конвекции-диффузии» (для трех компонент скорости, температуры, энергии турбулентности, влажности, дождевых капель, снежинок, облачной влаги, ледяных кристаллов, крупы) и несколько замыкающих алгебраических соотношений, вычисления в которых проводятся независимо по вертикальному координатному направлению.

Для определения наиболее перспективного способа построения эффективного параллельного численного метода решения уравнений модели TSUNM3 далее будем рассматривать только одно обобщенное дифференциальное уравнение конвективно-диффузионного переноса:

$$\begin{aligned} & \frac{\partial \rho \Phi}{\partial t} + \frac{\partial \rho U \Phi}{\partial x} + \frac{\partial \rho V \Phi}{\partial y} + \frac{\partial \rho W \Phi}{\partial z} = \\ & = \frac{\partial}{\partial x} \left(K_{xy} \frac{\partial \Phi}{\partial x} \right) + \frac{\partial}{\partial y} \left(K_{xy} \frac{\partial \Phi}{\partial y} \right) + \frac{\partial}{\partial z} \left(K_z \frac{\partial \Phi}{\partial z} \right) + S_{\Phi}, \end{aligned} \quad (1)$$

где t, x, y, z – время и координаты (ось Oz направлена вертикально вверх); U, V, W – компоненты вектора скорости, удовлетворяющего условию соленоидальности; ρ – плотность; K_{xy}, K_z – коэффициенты турбулентного обмена; S_{Φ} – источниковый член.

В качестве граничных условий для уравнения (1) на всех границах задаются простые условия Неймана (производная по нормали к поверхности равна нулю):

$$\frac{\partial \Phi}{\partial n} = 0. \quad (2)$$

В качестве начального условия используется нулевое значение концентрации примеси во всей области исследования:

$$\Phi(0, x, y, z) = 0. \quad (3)$$

Численный метод решения

Для дискретизации используется декартовая равномерная по горизонтальным направлениям и сгущающаяся к поверхности Земли сетка. Аппроксимация дифференциальной задачи (1) – (3) выполнена методом конечного объема со вторым порядком аппроксимации по времени и пространству. При аппроксимации конвективных членов уравнения (1) используются монотонизированные линейные противопотоковые схемы Ван Лира [5]. Для обеспечения второго порядка точности по времени используется комбинация неявной схемы Кранка – Николсон для вертикальной диффузии и явной схемы Адамса – Бэшфорда для остальных слагаемых уравнения (1):

$$\begin{aligned} & \frac{\rho^{n+1} \Phi_{i,j,k}^{n+1} - \rho^n \Phi_{i,j,k}^n}{\tau} Vol_{i,j,k} + \\ & + \frac{3}{2} \iiint_{Vol_{i,j,k}} \left[\frac{\partial \rho U \Phi}{\partial x} + \frac{\partial \rho V \Phi}{\partial y} + \frac{\partial \rho W \Phi}{\partial z} - \right. \\ & \left. - \frac{\partial}{\partial x} \left(K_{xy} \frac{\partial \Phi}{\partial x} \right) - \frac{\partial}{\partial y} \left(K_{xy} \frac{\partial \Phi}{\partial y} \right) - S_{\Phi} \right] dx dy dz - \end{aligned}$$

$$\begin{aligned}
& -\frac{1}{2} \iiint_{Vol_{i,j,k}} \left[\frac{\partial \rho U \Phi}{\partial x} + \frac{\partial \rho V \Phi}{\partial y} + \frac{\partial \rho W \Phi}{\partial z} - \right. \\
& \left. - \frac{\partial}{\partial x} \left(K_{xy} \frac{\partial \Phi}{\partial x} \right) - \frac{\partial}{\partial y} \left(K_{xy} \frac{\partial \Phi}{\partial y} \right) - S_{\Phi} \right]^{n-1} dx dy dz - \\
& -\frac{1}{2} \iiint_{Vol_{i,j,k}} \left[\frac{\partial}{\partial z} \left(K_z \frac{\partial \Phi}{\partial z} \right) \right]^{n+1} dx dy dz + \frac{1}{2} \iiint_{V_{i,j,k}} \left[\frac{\partial}{\partial z} \left(K_z \frac{\partial \Phi}{\partial z} \right) \right]^n dx dy dz = 0, \quad (4)
\end{aligned}$$

где n – номер слоя по времени; $Vol_{i,j,k}$ – значение конечного объема; здесь и далее нижние индексы i, j, k указывают, что переменная определяется в центре конечного объема.

Неявная аппроксимация для вертикального диффузионного переноса позволяет избежать жесткого ограничения на шаг интегрирования по времени. Результатом приближенного интегрирования является система линейных алгебраических уравнений с трехдиагональной матрицей следующего вида:

$$\begin{aligned}
& a_{i,j,k} \Phi_{i,j,k-1}^{n+1} - b_{i,j,k} \Phi_{i,j,k}^{n+1} + c_{i,j,k} \Phi_{i,j,k+1}^{n+1} = d_{i,j,k}; \\
& i = \overline{1, Nx}; \quad j = \overline{1, Ny}; \quad k = \overline{1, Nz}; \quad n = 1, 2, 3, \dots
\end{aligned} \quad (5)$$

Здесь $a_{i,j,k}$, $b_{i,j,k}$, $c_{i,j,k}$, $d_{i,j,k}$ – коэффициенты, вычисляющиеся явно в результате выполнения приближенного численного интегрирования уравнения (4). Выбранный способ аппроксимации по времени позволяет использовать при численном решении экономичный метод прогонки по вертикальной координате. Величина шага интегрирования по времени τ выбиралась из условия устойчивости разностной схемы.

Программная реализация численного алгоритма

Программная реализация алгоритма численного решения уравнения переноса выполнена на языке программирования C++. Ядро вычислительной части алгоритма содержит два цикла по i и j , перебирающих последовательно все узлы на плоскости xOy . Для каждого узла (i, j) выполняется отыскание численного решения в вертикальном направлении с использованием метода прогонки (рис. 1).

Оценка времени работы программной реализации выполнена для случая моделирования переноса примеси от постоянно действующего точечного источника в области с горизонтальными размерами (L)

50×50 км и высотой (H) 0,6 км. Источник примеси смещен относительно центра области исследования на $0,007L$ и $0,007L$ по осям Ox и Oy соответственно и расположен на высоте $0,0101H$. Компоненты скорости принимались постоянными: $U = 1$ м/с, $V = 1$ м/с, $W = 0$ м/с. Коэффициенты диффузии также принимались постоянными $K_{xy} = K_z = 100$ м²/с. Шаг по времени τ выбирался равным 6 с, а время окончания моделирования – 5000τ . Расчет проводился на сетке $256 \times 256 \times 32$.

```

1  For по времени
2
3  /* Циклы по пространственным координатам x и y */
4  For по i от 1 до Nx с шагом 1
5      For по j от 1 до Ny с шагом 1
6
7          /* Отыскание решения вдоль i-й, j-й
8             вертикальной линии методом прогонки */
9          For по k от 1 до Nz с шагом 1
10             Расчет коэффициентов a, b, c, d для
11             разностного уравнения (5). При вычислениях
12             используется решение с предыдущего шага
13             по времени  $\Phi_0 = \Phi^n$ ,  $\Phi_{00} = \Phi^{n-1}$ ;
14             Вычисление прогоночных коэффициентов
15             P[k], Q[k]
16
17             For по k от Nz до 1 с шагом -1
18                  $\Phi[i][j][k] \leftarrow P[k] * \Phi[i][j][k+1] + Q[k];$ 
19
20         For по i от 1 до Nx с шагом 1
21             For по j от 1 до Ny с шагом 1
22                 For по k от 1 до Nz с шагом 1
23                      $\Phi[i][j][k] \leftarrow \Phi_0 [i][j][k]$ 
24                      $\Phi_0[i][j][k] \leftarrow \Phi_{00}[i][j][k]$ 
25

```

Рис. 1. Упрощенное представление алгоритма численного решения рассматриваемой задачи

Для распараллеливания использовались технологии параллельного программирования Open Multi Processing (OpenMP), Message Passing Interface (MPI) и их комбинация. Вычисления проводились на кластере ТГУ Cyberia. Один вычислительный узел кластера имеет следующие характеристики: 48Gb RAM, 2xIntel® Xeon® X5670 (2,93GHz), 1xGPU NVIDIA Geforce RTX2080 Ti. Среднее время счета последовательной

программы на одном вычислительном узле кластера ТГУ Cyberia – 1915,9 с.

OpenMP – это библиотека для разработки параллельных программ, ориентированных на вычислительные системы с общей памятью. В ее состав входит набор директив компилятора, процедур и переменных окружения, которые предназначены для многопоточного программирования. Программирование осуществляется путем вставки директив компилятора в исходный код последовательной программы. OpenMP-программа может быть использована как последовательная, так как директивы OpenMP игнорируются последовательным компилятором. Для исполнения разработанного последовательного алгоритма в многопоточном режиме достаточно добавить перед вычислительными циклами по i, j директиву, указывающую на параллельное исполнение следующего за ней блока.

```
3      /* Циклы по пространственным координатам x и y */
      #pragma omp parallel for private (P, Q, j, k)
4      For по i от 1 до Nx с шагом 1
      #pragma omp parallel for private (P, Q, k)
5      For по j от 1 до Ny с шагом 1
```

При расчетах получено, что распараллеливание двух вложенных циклов дает такие же результаты, как и распараллеливание только внешнего цикла i . Использование ключа SCHEDULE с параметрами STATIC, DYNAMIC, GUIDED не дает ускорения по сравнению с распараллеливанием по умолчанию.

Параллельное программирование с использованием библиотеки OpenMP можно без преувеличения считать простым с точки зрения сложности применения. На рис. 2 представлено ускорение работы OpenMP-программы в зависимости от числа используемых процессов.

MPI – это библиотека передачи сообщений между параллельными процессами на системах с распределенной памятью. При этом функции библиотеки MPI реализованы и для систем с общей памятью, что дает возможность провести сравнение параллельных MPI- и OpenMP-реализаций в равных условиях (на одной вычислительной системе).

В отличие от OpenMP, использование MPI подразумевает самостоятельное распределение вычислительной работы по процессам и организацию обмена данными средствами библиотеки передачи сообщений MPI. Именно поэтому программисту предстоит больший объем работы и большая ответственность за качество ее выполнения.

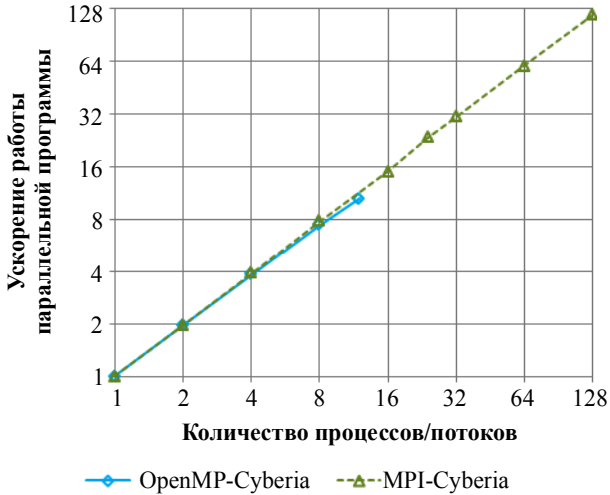


Рис. 2. Ускорение MPI (слева) и OpenMP (справа) программ

При использовании технологии MPI в качестве основного подхода распараллеливания выбрана двумерная (2D) декомпозиция сеточной области на подобласти. При разработке параллельной программы используется ряд функций MPI для создания двумерной декартовой логической топологии, создания пользовательских типов данных для обеспечения эффективной пересылки между процессами, а также функции `MPI_Isend()` и `MPI_Irecv()` для проведения парных неблокирующих пересылок значений сеточной функции в соседние подобласти двумерной XY-декомпозиции сеточной области.

Выбор в пользу двумерной декомпозиции сделан потому, что 2D-декомпозиция является более перспективной в сравнении с 1D-декомпозицией с точки зрения масштабирования на большое количество задействованных процессов [6]. В то же время 3D-декомпозиция не подходит для решения данной задачи, так как метод прогонки «плохо» поддается распараллеливанию.

Выбор двумерной геометрической декомпозиции подразумевает распределение каждому процессу, задействованному в вычислениях, некоторой сеточной подобласти со всеми неизвестными, принадлежащими ей. Для распределения подобластей по процессам функцией `MPI_Cart_create()` создается двумерная решетка процессов. Меняется диапазон изменения циклов i и j , теперь $i = 2 \cdot N_x 0 + 1$, $j = 2 \cdot N_y 1 + 1$, где

$Nx0 = Nx/dims[0]$, $Ny1 = Ny/dims[1]$ – это количество узлов горизонтальной сетки, которое будет обрабатываться каждым вычислительным процессом; $dims[0]$ и $dims[1]$ – количество процессов по координатам Ox и Oy ; $dims[0] * dims[1]$ – общее количество процессов, а $Nx0$ и $Ny1$.

В разностной схеме в горизонтальной плоскости используется девятиточечный шаблон крест, поэтому для расчета в приграничных узлах каждой области требуются данные с соседней подобласти. Здесь используется стандартный прием с созданием фиктивных ячеек [7].

На примере декомпозиции области исследования по оси Ox это реализуется следующим образом. На каждом процессе создаются фиктивные ячейки с номерами 0 и 1 слева и номерами $Nx0+2$ и $Nx0+3$ справа, ячейки с 2 по $Nx0+1$ являются внутренними. После каждой итерации по времени в фиктивные ячейки пересылаются данные. Сначала каждый процесс, кроме нулевого, пересылает содержимое ячеек с номером 2 и 3 налево и принимает в ячейки $Nx0+2$ и $Nx0+3$ значения справа (рис. 3). Аналогичным образом осуществляется пересылка данных в другом направлении.

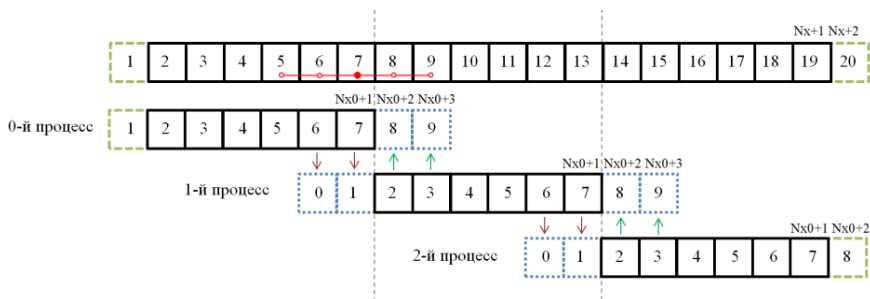


Рис. 3. Иллюстрация создания фиктивных ячеек и пересылки данных

Ускорение полученного алгоритма в зависимости от числа используемых процессов представлено на рис. 2. При проведении расчетов на кластере Cyberia получено практически линейное ускорение, а использование 128 ядер кластера Cyberia позволяет получить решение за 16,3 с. Сравнение MPI- и OpenMP-программ показывает близкое ускорение при использовании от 1 до 12 ядер.

Проведено исследование эффективности использования гибридного (двухуровневого) распараллеливания MPI+OpenMP, когда взаимодействие между вычислительными узлами осуществляется с использованием MPI, а распараллеливание внутри узла выполняется средствами

OpenMP. Для проведения эксперимента было выделено 12 двухпроцессорных узлов кластера ТГУ Cyberia (144 ядра).

При расчетах на каждом вычислительном узле всегда были задействованы все 12 ядер процессора, но менялось соотношение числа MPI-процессов/OpenMP-потоков следующим образом: 12/1, 6/2, 4/3, 3/4, 2/6, 1/12. Первая цифра указывает на количество MPI-процессов, которое приходится на вычислительный узел, вторая – на количество OpenMP-потоков, порождаемых в рамках одного MPI-процесса. Например, 12/1 – это 12 MPI-процессов по одному потоку на процесс (чистый MPI), а 1/12 – это 1 процесс с 12 OpenMP-потоками.

На рис. 4 представлено время выполнения программы при порождении разного количества потоков в рамках одного MPI-процесса. Наибольший эффект от гибридного распараллеливания получен при использовании 24 или 36 ядер. В этом случае полученное ускорение было на 16% больше, по сравнению с «чистым» MPI-распараллеливанием. При изменении числа задействованных узлов кластера от 4 до 10 (от 48 до 120 ядер) ускорение по сравнению с MPI-распараллеливанием выросло только на 7,5%. При использовании 144 ядер прирост ускорения составил 3,8%.

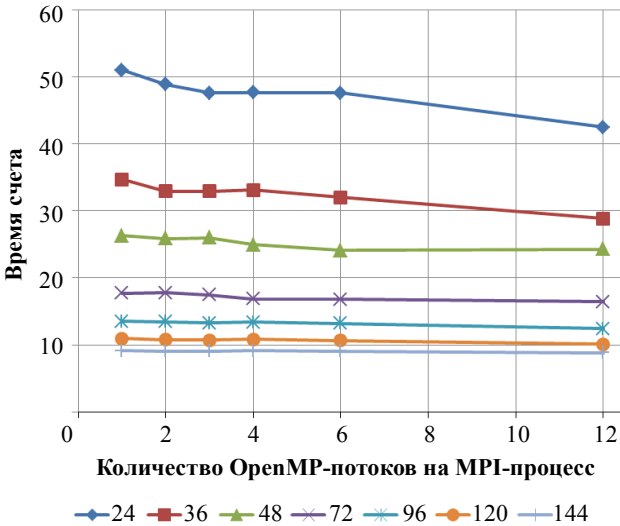


Рис. 4. Время выполнения программы при порождении разного количества потоков в рамках одного MPI-процесса

Таким образом, в рассматриваемой задаче получено, что наибольший эффект от гибридного распараллеливания проявляется при небольшом количестве задействованных ядер центрального процессора.

Если программный код будет содержать несколько программных модулей с большим объемом передаваемой между ними информации в процессе вычислений, что является свойством численной реализации математических моделей с большим количеством (более десяти) уравнений вида (1), то целесообразнее использовать технологию MPI.

Заключение

Результаты решения нестационарной 3D-задачи переноса примеси с помощью полунявной разностной схемы на структурированной сетке показали, что ускорение и эффективность OpenMP-программы практически не уступает MPI-программе при использовании небольшого (до 12) числа параллельных процессов. Кроме того, параллельная реализация программ с использованием технологии OpenMP является более простой, поэтому при решении вычислительных задач на компьютерах с общей памятью и небольшим числом процессоров/ядер предпочтительным является применение технологии распараллеливания OpenMP.

Проведение расчетов при использовании до 145 вычислительных ядер показало, что для системы с распределенной памятью MPI является более перспективной технологией распараллеливания.

СПИСОК ЛИТЕРАТУРЫ

1. *Starchenko A.V., Bart A.A., Kizhner L.I., Danilkin E.A.* Mesoscale meteorological model TSUNM3 for the study and forecast of meteorological parameters of the atmospheric surface layer over a major population center // Tomsk State University Journal of Mathematics and Mechanics. – 2020. – No. 66. – P. 35–55.
2. *Powers J.G., Klemp J.B., Skamarock W.C.* The weather research and forecasting model: Overview, system efforts and future directions // Bull. Amer. Meteor. Soc. – 2017. – V. 98. – P. 1717–1737.
3. *Ridwan R., Kistijantoro A.I., Kudsy M. & Gunawan D.* Performance evaluation of hybrid parallel computing for WRF model with CUDA and OpenMP // 3rd International Conference on Information and Communication Technology (ICoICT) (Nusa Dua). – 2015. – P. 425–430.
4. *Tolstykh M.A., Fadeev R.Yu., Shashkin V.V., Goyman G.S., Zaripov R.B., Kiktev D.B., Makhmorylova S.V., Mizyak V.G., Rogutov V.S.* Multiscale global atmosphere model SL-AV: the results of medium-range weather forecasts // Russ. Meteorol. Hydrol. – 2018. – V. 43. – P. 773–779.

5. *Van Leer B.* Towards the ultimate conservative difference scheme: II. Monotonicity and conservation combined in a second order scheme // *Journal of Computational Physics.* – 1974. – V. 14. – P. 361.
6. *Danilkin E.A., Starchenko A.V.* High performance computation for large eddy simulation // *LNCS series.* – 2010. – V. 6083. – P. 163.
7. *Starchenko A.V., Danilkin E.A., Semenova A., Bart A.A.* Parallel algorithms for a 3D photochemical model of pollutant transport in the atmosphere // *Communications in Computer and Information Science.* – 2016. – V. 687. – P. 158.

Старченко Александр Васильевич, докт. физ.-мат. наук, профессор, зав. кафедрой вычислительной математики и компьютерного моделирования НИ ТГУ; starch@math.tsu.ru;
Данилкин Евгений Александрович, канд. физ.-мат. наук, ст. науч. сотр. регионального научно-образовательного математического центра НИ ТГУ, ст. науч. сотр. лаборатории прогнозирования состояния атмосферы ИОА СО РАН; ugin@math.tsu.ru;
Проханов Сергей Анатольевич, программист I категории, ведущий программист НИ ТГУ; viking@math.tsu.ru;
Лещинский Дмитрий Викторович, мл. науч. сотр. регионального научно-образовательного математического центра НИ ТГУ, мл. науч. сотр. лаборатории прогнозирования состояния атмосферы ИОА СО РАН; 360flip182@gmail.com.

Параллельная реализация некоторых итерационных методов для решения системы линейных алгебраических уравнений в мезомасштабной метеорологической модели атмосферного пограничного слоя*

Е.А. Данилкин^{1,2}, А.В. Старченко¹,
Е.А. Каратаева^{1,2}, К.В. Юмин¹, И.В. Дель^{1,2},
Н.С. Смлян¹, Д.В. Лещинский^{1,2}

¹Томский государственный университет, г. Томск, Россия

²Институт оптики атмосферы им. В.Е. Зуева СО РАН, г. Томск, Россия

Представлены некоторые итоги выполнения проекта «Разработка параллельных вычислительных алгоритмов и программ для мезомасштабной модели численного прогноза погоды» в рамках ежегодно проводимой Международным математическим центром НГУ Большой математической мастерской [1]. Рассматриваются результаты исследований участников проекта по выбору эффективного параллельного алгоритма решения системы линейных алгебраических уравнений для мезомасштабной метеорологической модели атмосферного пограничного слоя.

Ключевые слова: MPI, OpenMP, параллельные вычисления, мезомасштабные метеорологические модели.

В настоящее время для краткосрочного прогнозирования и исследования погоды создаются и совершенствуются численные модели высокого пространственного разрешения, позволяющие предсказывать локальные региональные атмосферные явления для разных районов земного шара. На базе Томского государственного университета разработана и используется для моделирования и численного прогнозирования погоды мезомасштабная метеорологическая модель высокого разрешения TSUNM3 [2].

* Подготовка проекта и участие исследовательской группы в работе Большой математической мастерской выполнены при поддержке Министерства науки и высшего образования РФ (соглашение № 075-02-2021-1392). Внедрение результатов выполнения проекта в модель TSUNM3 и разработка нового метода расчета поправки давления выполнены за счет гранта Российского научного фонда (проект №19-71-20042).

Современные математические модели мезомасштабных (региональных) процессов в атмосфере включают систему нелинейных уравнений в частных производных, представляющих собой математическую запись основных законов сохранения. Решение такой системы с дополнительными начальными и граничными условиями можно получить только численно с применением суперкомпьютеров. Для этого строится сетка, покрывающая область исследования, дифференциальные уравнения заменяются конечно-разностными соотношениями, неизвестными в которых являются значения метеорологических параметров атмосферы в узлах сетки. Таким образом, решение системы дифференциальных уравнений сводится к решению нескольких больших систем линейных алгебраических уравнений (СЛАУ). В свою очередь, проблема решения СЛАУ с использованием методов вычислительной математики и компьютеров хорошо изучена, и численная реализация такого решения проще с точки зрения программирования.

При оперативном численном предсказании и моделировании атмосферных процессов необходимо в кратчайшие сроки представить результаты этого моделирования. Возможность ускорить процесс получения численного решения дает применение многопроцессорной и многоядерной вычислительной техники. Рабочая версия модели TSUNM3 распараллелена с использованием библиотеки MPI. Параллельная программа обеспечивает выполнение 24-часового прогноза за 17 минут процессорного времени на 144 ядрах вычислительного кластера ТГУ Cyberia.

В рамках выполнения заявленного проекта на Большой математической мастерской [1] перед исследовательской группой была поставлена задача реализовать несколько методов численного решения разностного аналога уравнения эллиптического типа для поправки давления и выбрать самый быстрый из них [3].

Математическая постановка задачи и численный метод ее решения

Математическая формулировка мезомасштабной модели TSUNM3 [2] включает одиннадцать нестационарных трехмерных неоднородных уравнений «конвекции-диффузии» (для трех компонент скорости, температуры, энергии турбулентности, влажности, дождевых капель, снежинок, облачной влаги, ледяных кристаллов, крупы), несколько замы-

кающих алгебраических соотношений (в том числе и для расчета радиационных потоков тепла в атмосфере) и процедуру расчета поправки к негидростатической части давления (решение разностного аналога эллиптического уравнения итерационным методом).

На рис. 1 представлено, в процентном соотношении, сколько времени занимают разные блоки программы при выполнении 24-часового прогноза. Как видно из диаграммы, расчет поправки давления занимает больше всего времени. Этот блок представляет собой нахождение численного решения разреженной СЛАУ с семидиагональной матрицей.

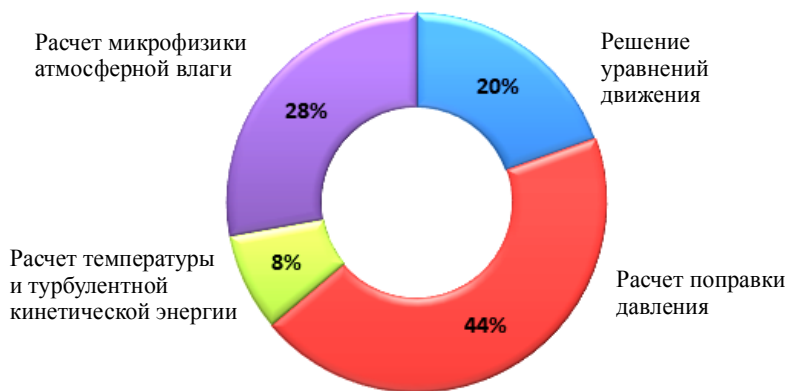


Рис. 1. Выполнение разных блоков программы TSUNM3

Процедура согласования полей скорости и давления

Для согласования конечно-разностных значений векторного поля скорости и давления на каждом шаге по времени используется схема «предиктор-корректор». Ее основная идея заключается в том, что сначала предсказываются значения компонент вектора скорости по известным на n -м слое по времени значениям сеточной функции давления p_h^n . Затем решается разностное уравнение эллиптического типа для поправки давления $p'_h = p_h^{n+1} - p_h^n$ и производится коррекция промежуточных полей скорости и давления. Эта операция осуществляется с требованием, чтобы скорректированные значения компонент скорости точно удовлетворяли разностному аналогу уравнения неразрывности [4].

Описанная схема включает в себя решение системы уравнений для поправки давления p'_h вида [2]

$$\begin{aligned}
 & ap_{i,j,k}(p')_{i,j,k} - ab_{i,j,k}(p')_{i,j,k-1} - at_{i,j,k}(p')_{i,j,k+1} - \\
 & - ae_{i,j,k}(p')_{i+1,j,k} - aw_{i,j,k}(p')_{i-1,j,k} - \\
 & - an_{i,j,k}(p')_{i,j+1,k} - as_{i,j,k}(p')_{i,j-1,k} = \\
 & = b_{i,j,k}(p'_h); i = \overline{1, Nx}; j = \overline{1, Ny}; k = \overline{1, Nz}.
 \end{aligned} \tag{1}$$

В рабочей версии модели для расчета поправки давления используется итерационный полинейный метод Гаусса – Зейделя с красным черным упорядочиванием узлов вычислительной сетки для каждого горизонтального уровня k и с неявным представлением искомым значений сеточной функции в узлах $(i, j, k+1)$, (i, j, k) , $(i, j, k-1)$ – т.е. вдоль вертикальных сеточных линий:

$$\begin{aligned}
 & -ab_{i,j,k}(p')_{i,j,k-1}^{l+1} + ap_{i,j,k}(p')_{i,j,k}^{l+1} - at_{i,j,k}(p')_{i,j,k+1}^{l+1} = \\
 & = ae_{i,j,k}(p')_{i+1,j,k}^l + an_{i,j,k}(p')_{i,j+1,k}^l + \\
 & + aw_{i,j,k}(p')_{i-1,j,k}^{l+1} + as_{i,j,k}(p')_{i,j-1,k}^{l+1} = b_{i,j,k}(p'_h); \\
 & i = \overline{1, Nx}; j = \overline{1, Ny}; k = \overline{1, Nz}.
 \end{aligned} \tag{2}$$

Работа в рамках Большой математической мастерской

Большая математическая мастерская – мероприятие, состоящее из двух недельных интенсивов (12 – 17 июля и 16 – 21 августа), разделенных четырехнедельным межмодулем [1].

Первый модуль

Работать с текущей версией модели TSUNM3 оказалось сложно по ряду причин. Программа написана на языке программирования FORTRAN, который оказался незнаком большинству участников проекта. Затратными по времени получились также компилирование, запуск и отладка всей модели целиком. Поэтому математическая постановка была упрощена до решения уравнения Пуассона в единичном кубе с граничными условиями первого рода:

$$\frac{\partial^2 P}{\partial x^2} + \frac{\partial^2 P}{\partial y^2} + \frac{\partial^2 P}{\partial z^2} = f(x, y, z);$$

$$(x, y, z) \in [0, 1] \times [0, 1] \times [0, 1];$$

$$P(x, y, z)|_G = \varphi(x, y, z). \quad (3)$$

Уравнение Пуассона является частным случаем эллиптического уравнения, решаемого в блоке расчета поправки давления в модели TSUNM3, и данное упрощение является распространенным подходом при проведении аналогичных исследований.

За первый модуль участники проекта познакомились с численными методами решения дифференциальных уравнений в целом и уравнения Пуассона в частности. С помощью метода конечного объема выполнили переход от дифференциального уравнения к системе линейных алгебраических уравнений. Вывели и выписали все необходимые формулы для вычисления коэффициентов СЛАУ и реализовали решение полученной системы методом Зейделя.

Межмодуль

На межмодуль каждый участник проекта получил задание реализовать один из перечисленных ниже методов решения СЛАУ: метод Зейделя; явный метод Булеева; метод сопряженных градиентов (CG); метод сопряженных градиентов с предобуславливанием методом Булеева (CG+B). Все участники успешно справились со своими задачами, и к началу второго модуля все перечисленные методы были реализованы программно.

Второй модуль

За время второго модуля участники проекта провели сравнение реализованных методов на примере численного решения уравнения Пуассона на равномерной декартовой сетке размером $50 \times 50 \times 50$. Расчеты выполнялись до достижения точности 10^{-10} по ошибке и невязке. В табл. 1 приведены две модельные задачи с известным аналитическим решением, на которых было выполнено сравнение методов решения СЛАУ. При сравнении анализировалось время работы программы и количество итераций, потребовавшееся для сходимости метода. Результаты проведенного исследования представлены в табл. 2, 3.

Таблица 1

Правая часть и аналитическое решение уравнения Пуассона для задач I и II

№ задачи	Аналитическое решение $P(x, y, z)$	Правая часть $f(x, y, z)$
I	$x \cdot (1-x) \cdot y \cdot (1-y) \cdot z \cdot (1-z)$	$-2 \cdot y \cdot z \cdot (1-y) \cdot (1-z) -$ $-2 \cdot x \cdot z \cdot (1-x) \cdot (1-z) -$ $-2 \cdot x \cdot y \cdot (1-x) \cdot (1-y)$
II	$x^2 \cdot y^2 \cdot z^2$	$6 \cdot x \cdot y^2 \cdot z^2 +$ $+2 \cdot x^3 \cdot z^2 + 2 \cdot x^3 \cdot y^2$

Таблица 2

Результаты сравнения методов решения СЛАУ для задачи I

Параметры	Метод Зейделя	Метод Булеева	Метод CG	Метод CG+B
Количество итераций	4623	1794	201	41
Время выполнения	45,84	28,64	1,77	1,93

Таблица 3

Результаты сравнения методов решения СЛАУ для задачи II

Параметры	Метод Зейделя	Метод Булеева	Метод CG	Метод CG+B
Количество итераций	4841	1862	238	64
Время выполнения	46,44	30,79	2,84	3,44

Результаты показали, что самым перспективным с точки зрения быстроедействия для рассматриваемых задач является метод сопряженных градиентов. Поэтому для метода сопряженных градиентов был разработан и программно реализован параллельный алгоритм решения рассматриваемой задачи.

Параллельная реализация метода сопряженных градиентов

К открытию метода сопряженных градиентов независимо пришли М. Хестенес и Э. Штифель. Он является наиболее предпочтительным по быстроедействию для симметричных положительно определенных систем [5]. Формулы классического метода сопряженных градиентов имеют следующий вид.

Задаются начальные данные:

$$x^0 - \text{начальное приближение; } r^0 = b - Ax^0; \quad p^0 = r^0.$$

Затем стартует итерационный процесс, который продолжается до тех пор, пока не будут выполнены условия сходимости:

$$v^m = Ap^m;$$

$$\alpha_m = \frac{(r^m, r^m)}{(p^m, v^m)};$$

$$x^{m+1} = x^m + \alpha_m p^m; \text{ если } x^m \text{ довольно точное, то остановка;}$$

$$r^{m+1} = r^m - \alpha_m v^m;$$

$$\beta_m = \frac{(r^{m+1}, r^{m+1})}{(r^m, r^m)};$$

$$p^{m+1} = r^{m+1} + \beta_m p^m.$$

Достаточным условием сходимости метода сопряженных градиентов являются симметричность и положительная определенность матрицы. Скорость сходимости метода сопряженных градиентов выше, чем скорость сходимости метода Зейделя.

Метод не требует никаких других операций кроме матрично-векторного произведения, вычисления скалярных произведений и сложения векторов. Все эти операции позволяют эффективно применять параллельные вычисления на компьютере с распределенной памятью. При распараллеливании использовались технология параллельного программирования MPI и геометрическая декомпозиция задачи. Геометрическая декомпозиция подразумевает разделение исходной области на равные части и распределение этих частей по ядрам вычислительной системы. В свою очередь, функции библиотеки MPI использовались для обмена сообщениями между параллельно выполнявшимися задачами.

В результате была построена и протестирована параллельная версия исходной последовательной программы и построен график ускорения работы параллельной программы в зависимости от количества задействованных процессов (рис. 2).

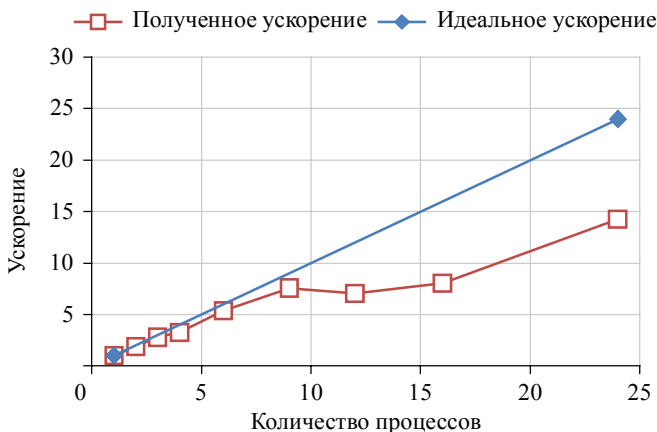


Рис. 2. Ускорение работы программной реализации метода сопряженных градиентов

Однако система линейных уравнений (1) не имеет строгого диагонального преобладания в силу наличия в преобразованной системе уравнений смешанных производных, которые аппроксимируются центрально-разностными формулами [2]. Поэтому использование популярного метода сопряженных градиентов в таких условиях не обеспечивает получения сходящегося итерационного процесса, и в связи с этим для внедрения в модель был выбран метод бисопряженных градиентов.

Разработка нового метода решения уравнения для поправки давления в модели TSUNM3

Стабилизирующий метод бисопряженных градиентов был разработан для решения систем линейных уравнений с несимметрической знакоопределенной матрицей, причем он избегает нерегулярных (с выбросами) промежуточных приближений «квадратичного» метода сопряженных градиентов [6]. Алгоритм метода BiCGStab реализуется следующим образом.

Задаются начальные данные:

x^0 – начальное приближение;

$r^0 = b - Ax^0$; $\rho_0 = \alpha = \omega_0 = 1$; $v^0 = p^0 = 0$;

\hat{r}^0 – произвольный вектор, для которого $(\hat{r}^0, r^0) \neq 0$,

например, $\hat{r}^0 = r^0$.

Затем стартует итерационный процесс, который продолжается до тех пор, пока не будут выполнены условия сходимости:

$$\rho_m = (\hat{r}^0, r^{m-1});$$

$$\beta_m = (\rho_m / \rho_{m-1})(\alpha / \omega_{m-1});$$

$$p^m = r^{m-1} + \beta_m (p^{m-1} - \omega_{m-1} v^{m-1});$$

$$v^m = Ap^m;$$

$$\alpha = \rho_m / (\hat{r}^0, v^m)$$

$$s = r^{m-1} - \alpha v^m;$$

$$t = As;$$

$$\omega_m = (t, s) / (t, t);$$

$x^m = x^{m-1} + \alpha p^m + \omega_m s$; если x^m довольно точное, то остановка;

$$r^i = s - \omega_i t.$$

Параллельная реализация метода бисопряженных градиентов

В качестве основного подхода распараллеливания используется двумерная (2D) декомпозиция сеточной области на подобласти. При разработке параллельной программы используется ряд функций MPI для создания двумерной декартовой логической топологии, создания пользовательских типов данных для обеспечения эффективной пересылки между процессами, а также функции MPI_Isend() и MPI_Irecv() для проведения парных неблокирующих пересылок значений сеточной функции в соседние подобласти двумерной XY-декомпозиции сеточной области [4].

После внедрения метода бисопряженных градиентов в модель TSUNM3 были проведены расчеты для оценки эффективности полученной параллельной реализации и сравнения с эффективностью полинейного метода Гаусса – Зейделя с красно-черным упорядочиванием.

Результаты расчетов показали (табл. 4), что метод бисопряженных градиентов при применении декомпозиции сеточной области и технологии передачи сообщений Message Passing Interface требует больше временных затрат с увеличением степени декомпозиции (количества разбиений) сеточной области.

Таблица 4

Результаты сравнения методов решения системы уравнений для поправки давления в модели TSUNM3

Метод решения СЛАУ	Последовательная программа		Параллельная программа	
	Общее время счета	Время расчета поправки давления	Общее время счета	Время расчета поправки давления
Метод Зейделя	1844,38	266,94	118,50	18,62
Метод бисопряженных градиентов	1770,75	187,51	166,88	69,36

Заключение

Участники работали в команде, обменивались опытом, совместно ставили задачи и преодолевали трудности. Команда проекта познакомилась с математическими аспектами задачи прогнозирования погоды и для решения предложенной задачи реализовала несколько методов решения СЛАУ.

Проведенное исследование также показало, что используемый в настоящее время в модели полинейный метод Зейделя продемонстрировал лучшие результаты при распараллеливании, поэтому в рабочей версии модели TSUNM3 он будет использоваться в качестве метода расчета поправки давления по умолчанию.

СПИСОК ЛИТЕРАТУРЫ

1. *Большая математическая мастерская* [Электронный ресурс]. – URL: <https://bmm.mca.nsu.ru/about>.
2. *Старченко А.В., Барт А.А., Кижнер Л.И., Данилкин Е.А.* Мезомасштабная метеорологическая модель TSUNM3 для исследования и прогнозирования состояния метеопараметров приземного слоя атмосферы над крупным населенным пунктом // *Вестн. Том. гос. ун-та. Математика и механика.* – 2020. – № 66. – С. 35–55.
3. *Разработка параллельных алгоритмов для мезомасштабной модели прогноза погоды* [Электронный ресурс]. – URL: <https://bmm.mca.nsu.ru/project/48>.

4. *Старченко А.В., Нутерман Р.Б., Данилкин Е.А.* Численное моделирование турбулентных течений и переноса примеси в уличных каньонах. – Томск: Изд-во Томского государственного университета, 2015. – 252 с.
5. *Крылов В.И., Бобков В.В., Монастырный П.И.* Вычислительные методы высшей математики: в 2 т. – Минск: Высшая школа, 1975. – Т. 2. – 670 с.
6. *Van der Vorst H.A.* Bi-CGSTAB: a fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems // *SIAM Journal*. – 1992. – V. 13(2). – P. 631–644.

Данилкин Евгений Александрович, канд. физ.-мат. наук, ст. науч. сотр. регионального научно-образовательного математического центра НИ ТГУ, ст. науч. сотр. лаборатории прогнозирования состояния атмосферы ИОА СО РАН; ugin@math.tsu.ru;

Старченко Александр Васильевич, докт. физ.-мат. наук, профессор, зав. кафедрой вычислительной математики и компьютерного моделирования НИ ТГУ; starch@math.tsu.ru;

Каратаева Екатерина Алексеевна, магистрант ММФ НИ ТГУ; nebosolnze@gmail.com;

Юмин Кирилл Викторович, студент ММФ НИ ТГУ; juminkir@mail.ru;

Дель Ирина Васильевна, магистрант ММФ НИ ТГУ; morgjerry@gmail.com;

Смиян Никита Сергеевич, аспирант ММФ НИ ТГУ; mr.turner.x@gmail.com;

Лещинский Дмитрий Викторович, мл. науч. сотр. регионального научно-образовательного математического центра НИ ТГУ, мл. науч. сотр. лаборатории прогнозирования состояния атмосферы ИОА СО РАН; 360flip182@gmail.com.

Повышение производительности LuNA-программ на основе воспроизведения трасс*

В.А. Перепёлкин, М.И. Иванов

*Институт вычислительной математики
и математической геофизики СО РАН, г. Новосибирск, Россия*

Трассировка (логгирование) исполнения распределенных параллельных программ является источником детальной информации о ходе вычислений, которая может быть использована для повышения эффективности программ (сокращение времени выполнения, расхода памяти и нагрузки на сеть). В системах автоматического конструирования параллельных программ трассировочная информация может быть собрана и использована автоматически. Рассматриваются средства системы автоматического конструирования параллельных программ LuNA, позволяющие осуществлять трассировку и на ее основе автоматически повышать эффективность конструируемых программ. Экспериментальное исследование на мультимпьютере показывает состоятельность данного подхода на примере задачи фильтрации двухфазной жидкости при наличии скважин.

Ключевые слова: *автоматическое конструирование параллельных программ, высокопроизводительные вычисления, распределенная память, система LuNA, трассировка программ, оптимизация параллельных программ.*

Разработка параллельных программ численного моделирования на мультимпьютерах¹ часто является сложной, трудоемкой и требующей специальной квалификации в области системного параллельного программирования. Это связано с необходимостью решать специфичные задачи разработки параллельных программ, такие как декомпозиция данных и вычислений, их распределение по вычислительным узлам, организация параллельной обработки распределенных данных, обеспечение доступа к требуемым данным на других вычислительных узлах и т.п. При этом необходимо обеспечить эффективность² параллельной

* Работа выполнена при поддержке бюджетного проекта ИВМиМГ СО РАН № 0251-2021-0005, а также при поддержке Российского научного фонда (проект № 19-11-00048).

¹ *Мультимпьютер* – параллельный вычислитель с распределенной памятью.

² *Эффективность* в работе понимается с точки зрения экономии ресурсов – времени выполнения, памяти, нагрузки на сеть.

программы, иначе целесообразность использования мультимпьютера для вычислений сомнительна. В связи с этим актуально создание и развитие существующих средств автоматизации конструирования параллельных программ, которые могут взять на себя решения некоторых или всех подобных задач. Главным вызовом для разработчиков таких средств в области высокопроизводительных вычислений является, как правило, обеспечение высокой (или хотя бы приемлемой) эффективности конструируемых программ. В общей постановке задача автоматического конструирования параллельной программы по некоторому высокоуровневому описанию алгоритма является алгоритмически трудно-решаемой, вследствие чего исследуются и развиваются различные частные и эвристические подходы, занимающие свои «ниши» в конструировании отдельных классов приложений и рассчитанные на отдельные классы вычислителей.

Ценным источником информации, которая может быть использована средствами автоматизации параллельного программирования, является трассировка (журналирование) исполнения параллельных программ. Идея трассировки близка к идее профилирования с той разницей, что при профилировании сохраняется, главным образом, статистическая информация о характеристиках исполнения программы (например, среднее время выполнения заданной подпрограммы). При трассировке же фиксируется полная информация о происходящих при исполнении событиях. Это позволяет использовать трассу (результат трассировки) для воспроизведения исполнения параллельной программы в точности так, как она была исполнена до этого (при трассировочном запуске). Главная выгода такого воспроизведения трассы состоит в том, что при воспроизведении трассы возможно исключить большую часть накладных расходов, связанных с работой исполнительской системы или других библиотек поддержки времени исполнения, относящихся к системе программирования. В особенности это касается интерпретируемых языков программирования, где накладные расходы на работу интерпретатора могут превышать собственно «полезные» вычисления.

В работе рассматривается задача трассировки и воспроизведения трасс в системе автоматического конструирования параллельных программ LuNA [1] на примере задачи фильтрации двухфазной жидкости при наличии скважин [2, 3].

Остальная часть статьи организована следующим образом. В разд. 1 излагаются ключевые особенности системы LuNA, необходимые для изложения основного материала. Разд. 2 посвящен вопросу реализации

идеи трассировки в системе LuNA. В разд. 3 приводятся результаты сравнительного тестирования производительности LuNA-программ при воспроизведении трасс и в обычном режиме на примере прикладной задачи.

1. Система LuNA

Система LuNA³ – это экспериментальная система автоматического конструирования параллельных программ численного моделирования на мультимпьютерах. Она разрабатывается в ИВМиМГ СО РАН (г. Новосибирск). Ее теоретический фундамент составляет теория структурного синтеза параллельных программ на вычислительных моделях [4].

В качестве входной программы система LuNA получает описание множества задач – фрагментов вычислений (ФВ), каждый из которых – триплет вида $\langle in, mod, out \rangle$. Здесь in и out – ссылочные выражения, обозначающие входные и выходные параметры триплета – фрагменты данных (ФД); а mod – это последовательная процедура без побочных эффектов, которая может быть вызвана для вычисления значений выходных ФД, если ей на вход подать значения входных, а собственно процесс вычислений рассматривается как исполнение таких триплетов по мере готовности их входных аргументов. По такому описанию система может автоматически конструировать и исполнять параллельную программу, автоматически распределяя ФВ и ФД по узлам мультимпьютера, динамически перераспределяя их при возникновении дисбаланса нагрузки, автоматически осуществляя передачу значений ФД по сети, осуществляя сборку мусора и выполняя прочие задачи, связанные с параллельным исполнением программы. Пользователь системы при этом освобождается от собственно параллельного программирования, ограничиваясь лишь описанием последовательного кода (модулей) и описанием множества триплетов на предметно-ориентированном языке LuNA. Это упрощает разработку, отладку и модификацию параллельных программ. С другой стороны, конструирование эффективной параллельной программы сложно, поэтому система LuNA накапливает и применяет различные эвристики и частные специализированные алгоритмы трансляции и исполнения LuNA-программ, что позволяет обеспечить достаточно эффективное конструирование параллельных про-

³ LuNA – Language for Numerical Algorithms.

грамм в частных случаях. Применение техники воспроизведения трасс призвано улучшить эффективность конструируемых программ и/или расширить класс приложений, для которых система LuNA способна конструировать программы приемлемой эффективности.

2. Воспроизведение трасс в системе LuNA

Какими бы сложными ни были алгоритмы трансляции и исполнения LuNA-программ в конечном счете исполнение сводится к запуску модулей с нужными параметрами (в соответствии с описанием триплетов). Каждый триплет запускается на одном из вычислительных узлов в определенный интервал времени. Поэтому для полного воспроизведения исполнения LuNA-программы достаточно зафиксировать при трассировке лишь один тип событий – начало исполнения ФВ. Для каждого события необходимо зафиксировать вычислительный узел, время, идентификаторы входных и выходных ФД и идентификатор модуля (процедуры).

Воспроизведение трассы при этом сводится к тому, чтобы на каждом узле мультимпьютера исполнять цикл следующего вида:

00. **WHILE** $|T| > 0$ ▷ Пока в трассе есть события
01. **POP** $c \leftarrow T$ ▷ Выбрать ФВ с минимальным
02. ▷ временем начала
03. **FOREACH** $d \leftarrow \text{ins}(c)$ ▷ для каждого входного ФД
04. **RECV** d ▷ дождаться получения его значения
05. ▷ по сети или локально
06. **END FOREACH**
07. **EXEC** c ▷ исполнить ФВ
08. **FOREACH** $d \leftarrow \text{outs}(c)$ ▷ для всех выходных ФД
09. **FOREACH** $e \leftarrow \text{cons}(d)$ ▷ для всех потребителей ФД d
10. **SEND** $e, \text{node}(e)$ ▷ отправить ФД e на узел
11. ▷ исполнения потребителя
12. **END FOREACH**
13. **END WHILE**

При воспроизведении трассы исключаются накладные расходы на работу исполнительской системы, задачами которой являются принятие решений о распределении и динамическом перераспределении ФВ и ФД по узлам, обеспечение доставки ФД с узлов их выработки на узлы

хранения, а оттуда – на узлы потребления, сборка мусора (удаление ненужных более ФД), мониторинг равномерности загрузки вычислительных узлов и ряд других функций. При воспроизведении трасс также существенно сокращается нагрузка на сеть, так как ФД передаются напрямую с узла производства на узел потребления, минуя все промежуточные пункты, имеющие место при нормальном исполнении LuNA-программы. Сборка мусора вырождается из сложного динамического механизма в тривиальный, так как все потребления ФД явно видны из трассы, и после последнего потребления на каждом узле, где есть копия ФД, эта копия удаляется сразу и без необходимости каких-либо коммуникаций.

Несмотря на преимущества воспроизведения трасс у этой техники есть существенные недостатки. Первая проблема состоит в том, что множество ФВ и ФД в общем случае зависит от данных. Например, если в программе имеется сходящийся за неизвестное заранее количество итераций процесс (цикл типа WHILE), то при разных входных данных может быть и разное количество итераций (а значит и ФВ и ФД). В этом случае воспроизведение трассы на других входных данных приведет к некорректной работе программы. Этот недостаток отчасти компенсируется тем, что во многих задачах множество ФВ и ФД не зависит от данных, а также тем, что если множество ФВ и ФД оказывается иным при воспроизведении трассы, то этот факт может быть выявлен автоматически путем сверки значений целочисленных ФД, фигурирующих в вычислении диапазонов циклов, условий в условных операторах или индексных выражениях. Если значения таких ФД при воспроизведении трассы оказываются отличными от их же значений, сохраненных при трассировке, то воспроизведение трассы останавливается, не производя некорректных расчетов.

Вторая проблема состоит в том, что в отличие от обычного исполнения LuNA-программы воспроизведение трасс не адаптивно. В трассе уже зафиксированы все решения о том, что, где и в каком порядке исполнять. Эти решения могут оказаться далекими от оптимальных, особенно если воспроизведение трасс выполняется на другом вычислителе, на других данных или в условиях иной внешней загрузки вычислителя. Отсутствие накладных расходов на работу исполнительной системы также сказывается. Этот недостаток может быть также отчасти преодолен. А именно, перед воспроизведением трасса может быть проанализирована и модифицирована. Например, изменены номера узлов, на ко-

торых выполняются те или иные ФВ. Это можно делать без риска внести ошибку в программу, т.к. процедуры, реализующие ФВ, не имеют побочных эффектов и, как следствие, могут выполняться на любых вычислительных узлах. Возможно также усложнение самой подсистемы воспроизведения трасс, например, добавлением поддержки динамической балансировки нагрузки на вычислитель.

3. Тестирование

Для тестирования была выбрана задача фильтрации двухфазной жидкости при наличии скважин [2, 3] как типичный представитель прикладных задач, требующих для проведения больших расчетов применения мультимикомпьютеров.

В статьях [2, 3] получены векторно-матричные формы систем для решения задачи фильтрации двухфазной жидкости при наличии скважин. Если для этих СЛАУ выписать задачу для дополнения Шура, то получим следующую матрично-векторную систему:

$$\left[B^T \mathbf{D}^{-1} B + \varepsilon C \right] \cdot \bar{u} = \bar{f}.$$

\mathbf{D} – симметрическая блочнодиагональная матрица с трехдиагональными блоками, $C = (\bar{e} \cdot \bar{e}^T)$ – симметрическая квадратная плотная матрица, но при этом умножение матрицы C на вектор осуществляется экономично: $C\bar{u} = [\bar{e}, \bar{u}] \bar{e}$, где $[\cdot, \cdot]$ – скалярное произведение в евклидовом пространстве. Матрица B является прямоугольной, количество строк матрицы превышает количество столбцов примерно в 3 раза, в каждой строке матрицы по 1–2 элемента (в основном по 2 элемента), а в каждом столбце по 3–6 элементов (в основном по 6 элементов).

В статье [2] доказывается положительная определенность матрицы $\left[B^T \mathbf{D}^{-1} B + \varepsilon C \right]$ для произвольного $\varepsilon > 0$, поэтому СЛАУ можно решать итерационным методом сопряженных градиентов. В статье [3] сказано, как искать ε , чтобы количество итераций метода сопряженных градиентов было близко к минимальному.

На каждом шаге итерационного процесса происходит умножение матрицы в квадратных скобках на вектор. Это является самой сложной операцией и отнимает львиную долю времени. Поэтому важно оптимизировать выполнение этой операции.

Умножение матрицы C на вектор может быть реализовано при помощи скалярного произведения векторов и операции умножения вектора на скаляр.

Умножение матрицы $B^T D^{-1} B$ на вектор состоит из умножения прямоугольной матрицы на вектор, решения СЛАУ с трехдиагональной матрицей методом прогонки и из умножения транспонированной прямоугольной матрицы на вектор.

Тестирование проводилось на мультикомпьютере «МВС 10п Торнадо» МСЦ РАН⁴, построенном на базе двухпроцессорных узлов с процессорами Intel Xeon E5-2690 и памятью 64 ГБ на узел. При тестировании использовались следующие значения параметров расчета: размер расчетной сетки $200 \times 200 \times 50$; количество шагов по времени 5; общее количество итераций взять из тестов; количество вычислительных узлов 4. Результаты измерения времени выполнения программы представлены в таблице.

Вариант программы	Время выполнения, с
LuNA	1541
Воспроизведение трасс	516
MPI-программа	204

По результатам измерений видно, что воспроизведение трасс позволило существенно сократить время выполнения программы по сравнению с обычным исполнением LuNA-программы, хотя разработанная вручную MPI-программа работает еще быстрее. Тем не менее времена сравнимы, отставание по производительности в несколько раз часто является приемлемым с учетом уменьшения времени на разработку и отладку параллельной программы по сравнению с ручным программированием. Так или иначе, воспроизведение трасс показало себя эффективной техникой повышения эффективности LuNA-программ на исследуемом приложении.

Заключение

Рассмотрена техника воспроизведения трасс и ее реализация в системе автоматического конструирования параллельных программ LuNA. Исследовано влияние этой техники на время выполнения программы на примере задачи фильтрации двухфазной жидкости при наличии сква-

⁴ <https://www.jscc.ru/resources/hpc/>

жин. Показано, что на данном приложении воспроизведение трасс позволило автоматически повысить эффективность конструируемой программы.

В дальнейшем планируется развивать технику воспроизведения трасс в соответствии с направлениями, обозначенными в разд. 2.

СПИСОК ЛИТЕРАТУРЫ

1. *Malyshkin V.E., Perepelkin V.A.* LuNA Fragmented programming system, Main functions and peculiarities of run-time subsystem // Parallel Computing Technologies. PaCT 2011. Lecture Notes in Computer Science. – V. 6873. – Berlin; Heidelberg: Springer, 2011. – https://doi.org/10.1007/978-3-642-23178-0_5.
2. *Ivanov M.I., Kremer I.A., and Laevsky Yu.M.* On the streamline upwind scheme of solution to the filtration problem // Siberian Electronic Mathematical Reports. – 2019. – V. 16. – P. 757–776.
3. *Ivanov M.I., Kremer I.A., and Laevsky Yu.M.* On wells modeling in filtration problems // Siberian Electronic Mathematical Reports. – 2019. – V. 16. – P. 1868–1884.
4. *Вальковский В.А., Малышкин В.Э.* Синтез параллельных программ и систем на вычислительных моделях. – Новосибирск: Наука, 1988. – 129 с.

Перепёлкин Владислав Александрович, науч. сотр. ИВМиМГ СО РАН;
perepelkin@ssd.sccc.ru;

Иванов Максим Игоревич, мл. науч. сотр. ИВМиМГ СО РАН; ivanov@sscc.ru.

Исследование возможностей языка программирования R для обработки больших данных*

И.А. Ботыгин^{1,2}, В.С. Шерстнёв², А.И. Шерстнёва²

¹Институт мониторинга климатических и экологических систем СО РАН, г. Томск, Россия

²Томский политехнический университет, г. Томск, Россия

Рассматривается использование пакетов языка программирования R для параллельного построения штрафных регрессионных моделей для больших наборов данных с ультразвуковых метеостанций, которые не могут быть сразу загружены в оперативную память компьютера. Файлы с наблюдениями хранятся на дисковом накопителе и считываются в память только во время подгонки модели. Матрицей ввода данных является `big.matrix`-объект. Создается функцией `as.big.matrix` из пакета `bigmemory`. Для распараллеливания вычислений используется стандарт OpenMP – программирование многопоточных приложений на многопроцессорных системах с общей памятью. Количество потоков для параллельных вычислений задается параметрически, исходя из имеющейся вычислительной мощности. Исследуемый инструментарий больше подходит для «широких» данных (с большим числом переменных-предикторов), чем для «длинных» данных (с большим количеством рядов наблюдений).

Ключевые слова: *регрессионный анализ, большие данные, язык R, параллельное программирование.*

В настоящее время анализ больших наборов данных и использование параллельной обработки являются активными направлениями исследований в области аналитики данных. Язык статистического программирования R, несмотря на внутреннюю реализацию, в основном, однопоточными операциями, может успешно использоваться вместе с различными подходами к параллельным вычислениям в сложных задачах анализа больших данных.

Параллелизм

Сразу отметим, что классическая многопоточная обработка, работающая в ядре операционной системы, недоступна на уровне R. Распа-

* Исследование выполнено при частичной финансовой поддержке РФФИ и Администрации Томской области в рамках научного проекта № 18-47-700005-а.

раллеливание в R происходит на более высоком уровне абстракции, который идентифицируется как уровень многопроцессного распараллеливания.

Ключевой технологией для параллельного выполнения кода на языке R является стандарт OpenMP – механизм написания параллельных программ для систем с общей памятью, в котором функционал генерирования эффективного параллельного кода возложен на компилятор. Свыше двухсот пакетов R, каждый для своих целей, используют директивы компилятора и библиотечные процедуры для формирования многопоточного кода.

Более явное распараллеливание можно осуществить при помощи пакета snow (Simple Network of Workstations), реализующего простой механизм использования набора рабочих станций для параллельных вычислений [1]. По сути, пакет snow – это инструмент для разработки кластерных приложений, предоставляющий общий высокоуровневый интерфейс к механизму построения программ по модели обмена сообщениями MPI (Message Passing Interface), к механизму передачи сообщений для библиотеки PVM (Parallel Virtual Machine), к сырым сокетам Raw sockets, осуществляющим прямую отправку и получение IP-пакетов без какого-либо форматирования транспортного уровня.

Упрощает параллельное программирование с использованием кластеров пакет snowfall (Easier cluster computing), построенный на основе пакета snow с использованием его сетевых и кластерных возможностей [2]. В частности, пакет snowfall предлагает дополнительную поддержку неявного параллельного выполнения, дополнительные вычислительные функции, расширенную обработку ошибок и множество функций для более удобного программирования. Инструментарий snowfall не добавляет уровень абстракции к базовому пакету snow, а рассматривается как дополнение для пользователей, ищущих большее удобство в использовании параллельных вычислений.

Другим расширением пакета snow, обеспечивающим поддержку простого в использовании параллельного программирования и отказоустойчивые и воспроизводимые приложения, является пакет snowFT (Fault Tolerant Simple Network of Workstations) [3].

Альтернативный вариант использования нескольких ядер на одной машине напрямую для процессно-параллельной работы обеспечивает пакет parallelLogger [4]. Пакет включает функции для параллельных вычислений с индикатором выполнения и возможностью остановки или продолжения при ошибках, функции ведения журнала, включая автома-

тическое ведение журнала ошибок и предупреждений, функции, используемые для автоматизации анализа. Поддерживается автоматизация вызова функций с отложенным выполнением (например, для параллельного выполнения функций).

В заключение небольшого обзора инструментария языка статистического программирования R для организации параллельных вычислений отметим пакет `future` (Unified Parallel and Distributed Processing in R for Everyone) [5]. По сути, пакет `future` представляет собой прикладной уровень (фактически, фреймворк) для последовательной и параллельной обработки выражений R. Единый унифицированный интерфейс прикладного программирования (Future API) обеспечивает возможность конструировать как простые, так и многоядерные, мультисессионные и кластерные структуры для параллельных вычислений. Одним из ключевых элементов пакета `future` является возможность реализации асинхронных вычислений.

Расширения пакета `future` [6–9] реализуют дополнительные бэкенды и механизмы абстракции для формирования крупномасштабных компьютерных экспериментов с помощью языка R.

Данные

Другим аспектом организации высокопроизводительных вычислений является проблема «больших данных». Базовыми инструментами в этой области являются платформа Apache Hadoop [10] и движок Apache Spark [11]. В языке R существует целая экосистема соответствующих сервисов для крупномасштабной обработки данных [12–20]. Особо стоит выделить пакет `dplyr` (A Grammar of Data Manipulation), для которого `sparkdplyr` может выступать в качестве бэкенда и соединения со Spark [21].

Для проведения программных экспериментов использовались данные, полученные с высокочастотной ультразвуковой метеостанции, расположенной на полигоне ИМКЭС СО РАН. Интервал между измерениями параметров – 12,5 мс. Суточный объем измеряемых данных только с одной станции превышает 90 МБ. Результаты измерений записываются на сервер хранения модулями с десятиминутными наблюдениями. Форматы файлов хранения разработаны с учетом конкретных характеристик данных и обеспечивают как общие, так и специфические варианты их использования.

Обработка таких больших данных дополнительно нагружает систему хранения. Даже извлечение для обработки по критериям запросов требуемых подмножеств данных из файлов, хранящихся на диске, требует значительного времени. Загружаются небольшие порции данных с диска в структуры данных языка в оперативной памяти, выполняется только частичный анализ, затем осуществляется переход к другим подмножествам данных и продолжение обработки и т.д. Но в условиях ограниченных ресурсов – это рациональный и стабильный способ для крупномасштабной обработки данных. Подход, основанный на отображении файлов в память (memory-mapped file), при недостатке ресурсов, наоборот, может быть значительно медленнее, чем стандартный файловый ввод-вывод.

Особенность выборки данных заключалась в том, что исследователям требовалось форматировать результаты измерений по интервалу усреднения и по сроку наблюдения. На рис. 1 отображено начальное окно модуля формирования выбора параметров.

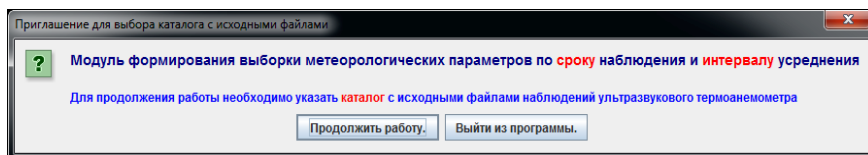


Рис. 1. Начальное окно модуля выборки

Для облегчения выбора срока наблюдений был реализован интерфейс в привычном для исследователя календарном виде (рис. 2).



Рис. 2. Окно установки срока наблюдений

Последним этапом выбора параметров является выбор интервала усреднения. Для выбора интервала усреднения необходимо указать время, за которое требуется усреднить данные. Для облегчения выбора времени были представлены таблицы, с помощью которых не требуется вводить данные, а возможен визуальный выбор механическим манипулятором (рис. 3).



Рис. 3. Выбор интервала усреднения

После выбора параметров усреднения появляется окно, в котором отображены сводные данные по сформированным параметрам (рис. 4).

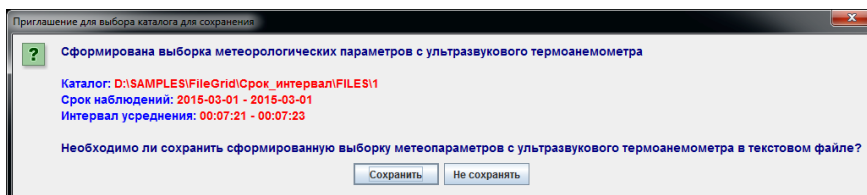


Рис. 4. Окно сводных данных по выборке

После нажатия кнопки «Сохранить» появляется новое окно, в котором осуществляется выбор каталога сохранения и выбор названия нового файла (рис. 5).

Данные сохраняются в текстовый файл, в котором отображены все данные о выборе срока и интервала усреднения, а также отображены данные о потраченном времени работы программы и таблица с данными (рис. 6).

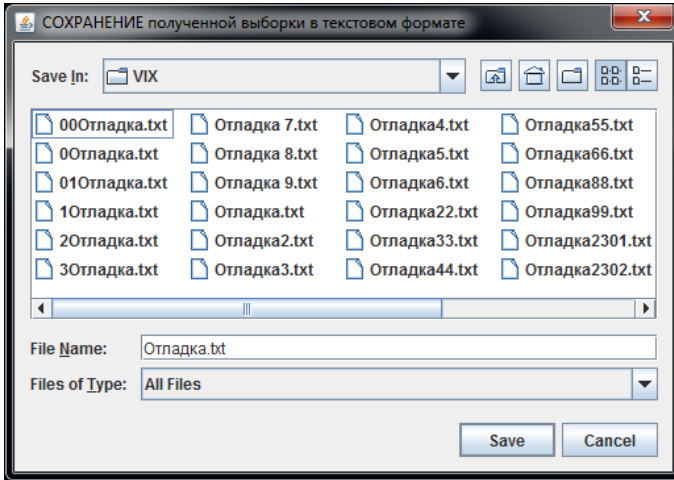


Рис. 5. Окно сохранения данных

Выборка метеорологических параметров <наблюдения с ультразвуковых термоанемометров>

каталог: D:\SAMPLES\FileId\срок_интервал\FILES\1

срок наблюдений: с 2015-03-01 по 2015-03-01

интервал усреднения: с 00:07:21 по 00:07:23

время обработки запроса = 191 мс

номер	дата и время	Температура	Ю. комп. ветра	В. комп. ветра	Верт. комп. ветра	давление	Влажность	ошибка
п/п	год-месяц-день час:мин:сек,мс	Град. Цельсия	м/сек	м/сек	м/сек	мм рт.ст.	%	код
1.	2015-03-01 00:07:21,001	-3.43	-2.93	0.28	-0.92	749.8	90.8	0.0
2.	2015-03-01 00:07:21,014	-3.43	-2.92	0.35	-0.92	749.8	90.79	0.0
3.	2015-03-01 00:07:21,026	-3.35	-3.09	0.25	-0.89	749.8	90.8	0.0
4.	2015-03-01 00:07:21,039	-3.4	-2.73	0.42	-0.92	749.8	90.8	0.0
5.	2015-03-01 00:07:21,051	-3.35	-2.74	0.43	-0.87	749.8	90.8	0.0
6.	2015-03-01 00:07:21,064	-3.39	-2.73	0.33	-0.86	749.8	90.8	0.0
7.	2015-03-01 00:07:21,076	-3.4	-2.8	0.4	-0.84	749.8	90.8	0.0
8.	2015-03-01 00:07:21,089	-3.4	-2.6	0.42	-0.94	749.8	90.81	0.0
9.	2015-03-01 00:07:21,101	-3.34	-2.81	0.35	-0.9	749.8	90.81	0.0
10.	2015-03-01 00:07:21,114	-3.41	-2.93	0.3	-0.88	749.8	90.81	0.0
11.	2015-03-01 00:07:21,126	-3.41	-2.89	0.28	-0.9	749.8	90.82	0.0
12.	2015-03-01 00:07:21,138	-3.39	-2.8	0.28	-0.87	749.8	90.82	0.0
13.	2015-03-01 00:07:21,151	-3.47	-2.84	0.23	-0.86	749.8	90.82	0.0
14.	2015-03-01 00:07:21,163	-3.45	-2.6	0.44	-1.0	749.8	90.82	0.0
15.	2015-03-01 00:07:21,176	-3.43	-2.63	0.25	-0.86	749.8	90.82	0.0
16.	2015-03-01 00:07:21,188	-3.32	-2.72	0.34	-0.93	749.8	90.82	0.0
17.	2015-03-01 00:07:21,201	-3.37	-2.8	0.18	-0.94	749.8	90.82	0.0
18.	2015-03-01 00:07:21,213	-3.33	-2.76	0.27	-1.0	749.7	90.82	0.0
19.	2015-03-01 00:07:21,226	-3.38	-2.64	0.36	-1.03	749.7	90.82	0.0
20.	2015-03-01 00:07:21,238	-3.41	-2.58	0.39	-1.05	749.8	90.81	0.0
21.	2015-03-01 00:07:21,251	-3.37	-2.52	0.51	-0.99	749.7	90.81	0.0
22.	2015-03-01 00:07:21,263	-3.36	-2.39	0.39	-0.94	749.7	90.81	0.0
23.	2015-03-01 00:07:21,276	-3.37	-2.52	0.42	-0.97	749.7	90.81	0.0
24.	2015-03-01 00:07:21,288	-3.34	-2.53	0.54	-1.0	749.7	90.81	0.0
25.	2015-03-01 00:07:21,301	-3.41	-2.58	0.36	-1.01	749.7	90.8	0.0
26.	2015-03-01 00:07:21,313	-3.44	-2.64	0.32	-1.0	749.7	90.79	0.0
27.	2015-03-01 00:07:21,326	-3.41	-2.59	0.39	-0.98	749.7	90.8	0.0
28.	2015-03-01 00:07:21,338	-3.47	-2.6	0.42	-1.0	749.7	90.8	0.0
29.	2015-03-01 00:07:21,350	-3.32	-2.6	0.25	-0.87	749.7	90.8	0.0
30.	2015-03-01 00:07:21,363	-3.44	-2.69	0.34	-0.99	749.7	90.8	0.0
31.	2015-03-01 00:07:21,375	-3.42	-2.6	0.31	-1.03	749.7	90.8	0.0
32.	2015-03-01 00:07:21,388	-3.38	-2.86	0.12	-1.0	749.7	90.8	0.0
33.	2015-03-01 00:07:21,400	-3.4	-2.91	0.07	-1.0	749.7	90.8	0.0
34.	2015-03-01 00:07:21,413	-3.4	-2.89	0.05	-0.95	749.7	90.8	0.0
35.	2015-03-01 00:07:21,425	-3.4	-2.83	0.1	-0.9	749.7	90.79	0.0
36.	2015-03-01 00:07:21,438	-3.33	-2.55	0.24	-0.93	749.7	90.78	0.0
37.	2015-03-01 00:07:21,450	-3.34	-2.73	0.19	-0.97	749.7	90.78	0.0
38.	2015-03-01 00:07:21,463	-3.43	-2.84	0.2	-0.97	749.7	90.79	0.0

Рис. 6. Таблица данных

Все программное обеспечение для выборки необходимых результатов наблюдений из файлового хранилища реализовано на языке объектно-ориентированного программирования Java.

Обработка

Одним из подходов для моделирования и анализа отношений между переменными является регрессионный анализ. Вместе с тем базовая модель в виде линейной регрессии из-за ограниченных возможностей часто неприменима к реальным данным. Используются вариации линейной регрессии – регрессии с функциями штрафа.

В экспериментах с $L1$ и $L2$ регуляризацией в качестве регрессоров для температуры воздуха были выбраны скорость горизонтального ветра, направление горизонтального ветра, атмосферное давление, относительная влажность.

Регрессия LASSO (лассо) использует регуляризацию $L1$ – взвешивает ошибки по их абсолютному значению (рис. 7). λ – фактор, контролирующий размер штрафа.

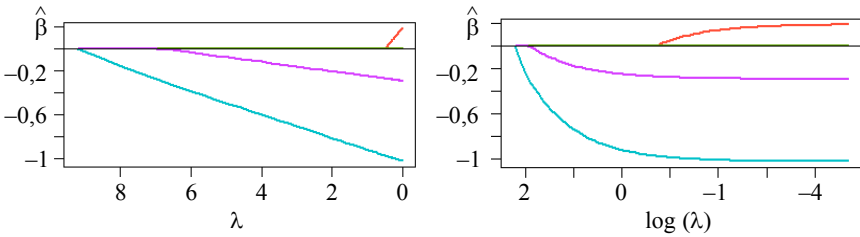


Рис. 7. Штраф Лассо

$L1$ -регуляризация часто приводит к разреженным моделям, так как некоторые коэффициенты могут стать нулевыми и, следовательно, не будут учитываться в составе модели.

Ridge regression (гребневая регрессия, ридж-регрессия) использует регуляризацию $L2$, которая взвешивает ошибки по их квадрату (рис. 8).

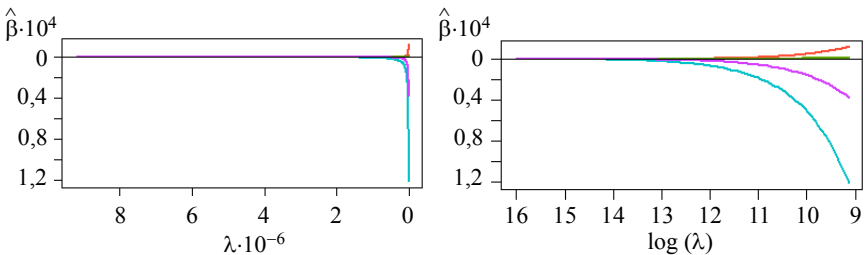


Рис. 8. Гребневая регрессия

При этом из-за природы регуляризации $L2$ ни один из коэффициентов при гребневой регрессии не становится нулевым – коэффициенты все больше приближаются к нулю.

Регрессия ElasticNet (эластичная сеть) – метод регрессии, в котором линейно объединяются штрафы регуляризаций $L1$ и $L2$ (рис. 9).

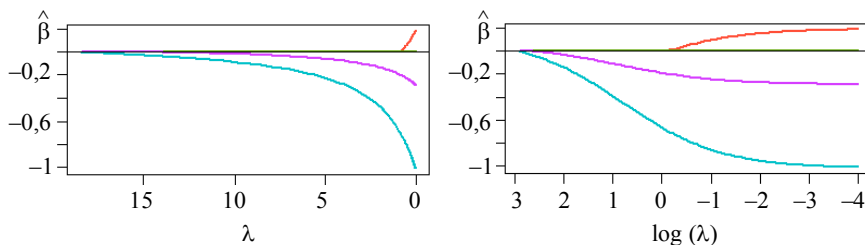


Рис. 9. Эластичная сеть

Заметим, что в регрессии ElasticNet возможно варьировать меру «смешанности» регуляризаций $L1$ и $L2$, что может привести либо к чисто гребневой регрессии, либо к чистой регрессии лассо.

Заключение

Приведены некоторые практические результаты при подгонке регрессионных моделей, штрафующих по методу лассо, гребня и эластичной сети для больших данных, которые не могут быть загружены в память.

Рассмотренный инструментарий языка статистического моделирования R позволяет исследователям при решении сложных задач анализа работать с большими файлами временных рядов и осуществлять многопроцессное распараллеливание.

СПИСОК ЛИТЕРАТУРЫ

1. *Snow*: Simple Network of Workstations [Электронный ресурс]. – URL: <https://cran.r-project.org/web/packages/snow/index.html>website (дата обращения: 12.04.2021).
2. *Snowfall*: Easier cluster computing (based on snow) [Электронный ресурс]. – URL: <https://cran.r-project.org/web/packages/snowfall/index.html> (дата обращения: 12.04.2021).
3. *SnowFT*: Fault Tolerant Simple Network of Workstations [Электронный ресурс]. – URL: <https://cran.r-project.org/web/packages/snowFT/index.html> (дата обращения: 12.04.2021).

4. *ParallelLogger*: Support for Parallel Computation, Logging, and Function Automation [Электронный ресурс]. – URL: <https://cran.r-project.org/web/packages/ParallelLogger/index.html> (дата обращения: 12.04.2021).
5. *Future*: Unified Parallel and Distributed Processing in R for Everyone [Электронный ресурс]. – URL: <https://cran.r-project.org/web/packages/future/index.html> (дата обращения: 22.05.2021).
6. *Future.batchtools*. A Future API for Parallel and Distributed Processing using 'batchtools' [Электронный ресурс]. – URL: <https://cran.r-project.org/web/packages/future.batchtools/index.html> (дата обращения: 22.05.2021).
7. *Future.apply*. Apply Function to Elements in Parallel using Futures [Электронный ресурс]. – URL: <https://cran.r-project.org/web/packages/future.apply/index.html> (дата обращения: 22.05.2021).
8. *Future.callr*. A Future API for Parallel Processing using 'callr' [Электронный ресурс]. – URL: <https://cran.r-project.org/web/packages/future.callr/index.html> (дата обращения: 22.05.2021).
9. *Future.tests*. Test Suite for 'Future' API Backends [Электронный ресурс]. – URL: <https://cran.r-project.org/web/packages/future.tests/index.html> (дата обращения: 22.05.2021).
10. *Apache Hadoop* [Электронный ресурс]. – URL: <https://hadoop.apache.org/> (дата обращения: 22.05.2021).
11. *Apache Spark* [Электронный ресурс]. – URL: <https://spark.apache.org/> (дата обращения: 22.05.2021).
12. *Spark.sas7bdat*. Read in 'SAS' Data into 'Apache Spark' [Электронный ресурс]. – URL: <https://cran.r-project.org/web/packages/spark.sas7bdat/index.html> (дата обращения: 22.05.2021).
13. *Sparkavro*. Load Avro file into 'Apache Spark' [Электронный ресурс]. – URL: <https://cran.r-project.org/web/packages/sparkavro/index.html> (дата обращения: 22.05.2021).
14. *Sparkbq*. Google 'BigQuery' Support for 'sparklyr' [Электронный ресурс]. – URL: <https://cran.r-project.org/web/packages/sparkbq/index.html> (дата обращения: 22.05.2021).
15. *Sparklyr*. R Interface to Apache Spark [Электронный ресурс]. – URL: <https://cran.r-project.org/web/packages/sparklyr/index.html> (дата обращения: 22.05.2021).
16. *Sparklyr.flint*. Sparklyr Extension for 'Flint' [Электронный ресурс]. – URL: <https://cran.r-project.org/web/packages/sparklyr.flint/index.html> (дата обращения: 22.05.2021).
17. *Sparklyr.nested*. A 'sparklyr' Extension for Nested Data [Электронный ресурс]. – URL: <https://cran.r-project.org/web/packages/sparklyr.nested/index.html> (дата обращения: 22.05.2021).
18. *Sparktf*. Interface for 'TensorFlow' 'TFRecord' Files with 'Apache Spark' [Электронный ресурс]. – URL: <https://cran.r-project.org/web/packages/sparktf/index.html> (дата обращения: 22.05.2021).
19. *Sparkwarc*. Load WARC Files into Apache Spark [Электронный ресурс]. – режим доступа: <https://cran.r-project.org/web/packages/sparkwarc/index.html> (дата обращения: 22.05.2021).

20. *Sparkxgb*. Interface for 'XGBoost' on 'Apache Spark' [Электронный ресурс]. – URL: <https://cran.r-project.org/web/packages/sparkxgb/index.html> (дата обращения: 22.05.2021).
21. *Dplyr*. A Grammar of Data Manipulation [Электронный ресурс]. – URL: <https://cran.r-project.org/web/packages/dplyr/index.html> (дата обращения: 22.05.2021).

Ботыгин Игорь Александрович, канд. техн. наук, доцент, ведущий программист лаборатории биоинформационных технологий ИМКЭС СО РАН, доцент Инженерной школы информационных технологий и робототехники ТПУ; bia@tpu.ru;

Шерстнёв Владислав Станиславович, канд. техн. наук, доцент, зав. кафедрой – руководитель отделения информационных технологий на правах кафедры Инженерной школы информационных технологий и робототехники ТПУ; vss@tpu.ru;

Шерстнёва Анна Игоревна, канд. физ.-мат. наук, доцент, доцент Школы базовой инженерной подготовки ТПУ; sherstneva@tpu.ru.

Influence of the pollution source location on the concentration of the impurity in street canyon*

D.V. Leshchinskiy, E.A. Danilkin, A.V. Starchenko

Tomsk State University, Tomsk, Russia

The work is devoted to the study of the flow structure and the nature of the impurities distribution in a street canyon, depending on the location of a constant intensity pollution source. The paper considers three cases when the source of the impurity is located: in the center at the bottom of the street canyon, at the bottom under leeward wall and at the bottom under windward wall.

Keywords: *street canyon, mathematical modeling, turbulent flow, impurity concentration.*

This work is devoted to the construction and testing of a mathematical model for studying the processes of transport of impurities in street canyons. Studies of this kind are relevant for understanding the microclimate of cities, are useful in planning new buildings and assessing areas of increased environmental hazard. Based on the developed original micro-scale mathematical model M2U [1], a study was carried out of the influence of the location of the pollution source on the flow structure and the nature of the distribution of impurities inside the street canyon. The mathematical model is based on the Reynolds-averaged Navier-Stokes equations. The system of differential equations is closed using a two-parameter model, which takes into account the influence of buoyancy forces on the turbulent flow structure, and the Boussinesq gradient-diffusion hypothesis. The numerical solution is carried out based on the finite volume method, a monotonized scheme for approximating convective terms, and the SIMPLE algorithm for matching the velocity and pressure fields [2].

In this work, the emphasis is on the study of changes in the maximum and average concentration in the breathing zone of a person (up to 2 m from the ground) and in the canyon as a whole. Three cases of the location of the pollution source were considered: in the center at the bottom of a street canyon, at the bottom at the leeward wall, and at the bottom at the windward wall.

* This work was supported by the Ministry of Science and Higher Education of Russia (agreement No. 075-02-2020-1479/1).

This study will help to understand how strongly the pollution source location will affect the accumulation of impurities in the street canyon. Thus, the data obtained can be used in the design of residential areas and parking lots.

Physical and mathematical formulation of the problem

Stationary isothermal turbulent air movement in the surface layer over an inhomogeneous underlying surface with elements of large-scale roughness is considered. Roughness elements are fixed, flow-tight obstacles (buildings). A point source of harmful emissions of constant intensity (vehicles) is considered. The dimensions of buildings are commensurate with the dimensions of the study area (Figure 1).

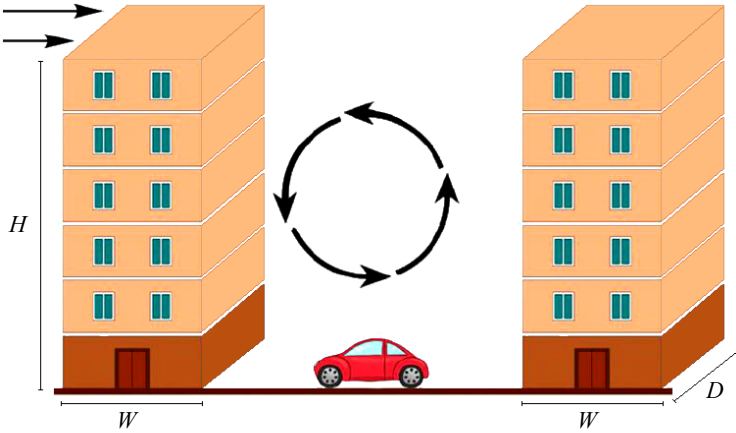


Fig. 1. Physical statement of the problem

The mathematical model includes Reynolds-averaged continuity equations, Navier-Stokes equations, impurity transport [2, 3] and heat transfer:

$$\frac{\partial \bar{u}_i}{\partial x_i} = 0; \tag{1}$$

$$\frac{\partial \bar{u}_i \bar{u}_j}{\partial x_j} = -\frac{1}{\rho} \frac{\partial \bar{p}}{\partial x_i} + \frac{\partial}{\partial x_j} \left(\nu \frac{\partial \bar{u}_i}{\partial x_j} \right) - \frac{\partial \overline{u'_i u'_j}}{\partial x_j}; \tag{2}$$

$$\frac{\partial \bar{u}_j \bar{C}}{\partial x_j} = \frac{\partial}{\partial x_j} \left(D \frac{\partial \bar{C}}{\partial x_j} \right) - \frac{\partial \overline{u'_j C'}}{\partial x_j} + S; \quad i, j = 1, 2, 3. \tag{3}$$

Here \bar{u}_i , u'_i – averaged and ripple projections of the velocity vector on the coordinate axis, \bar{p} – pressure, ρ – density, ν – kinematic viscosity of air, \bar{C} – averaged impurity concentration, S – function describing the distribution of impurity sources, $\overline{u'_i u'_j}$ – Reynolds stress tensor, $D = \nu/Sc$ – diffusion coefficient, Sc – Schmidt number.

The closure of the system of equations (1) – (3) is carried out using the two-parameter « k - ε » model and the Boussinesq gradient-diffusion hypothesis [2, 4]:

$$\overline{u'_i u'_j} = -\nu_T \left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right) + \frac{2}{3} k \delta_{ij}; \quad \overline{C'_i u'_j} = -\frac{\nu_T}{Sc_T} \frac{\partial \bar{C}}{\partial x_j}; \quad (4)$$

$$\nu_T = C_\mu \frac{k^2}{\varepsilon}; \quad P = -\overline{u'_i u'_j} \frac{\partial \bar{u}_i}{\partial x_j}; \quad (5)$$

$$\frac{\partial \bar{u}_j k}{\partial x_j} = \frac{\partial}{\partial x_j} \left(\left(\nu + \frac{\nu_T}{\sigma_k} \right) \frac{\partial k}{\partial x_j} \right) + P - \varepsilon; \quad (6)$$

$$\frac{\partial \bar{u}_j \varepsilon}{\partial x_j} = \frac{\partial}{\partial x_j} \left(\left(\nu + \frac{\nu_T}{\sigma_\varepsilon} \right) \frac{\partial \varepsilon}{\partial x_j} \right) + \frac{\varepsilon}{k} (C_{\varepsilon 1} P - C_{\varepsilon 2} \varepsilon), \quad (7)$$

where ν_T – turbulent viscosity, k – kinetic energy of turbulence, ε – dissipation of turbulent kinetic energy. Turbulence model coefficients: $\sigma_k = 1.0$, $\sigma_\varepsilon = 1.3$, $C_{\varepsilon 1} = 1.44$, $C_{\varepsilon 2} = 1.92$, $C_\mu = 0.09$, $Sc_T = 0.5$ – turbulent Schmidt number.

A full description of the mathematical model used and the numerical method for its solution is given in [5].

Simulation of flow and parametric calculations

The analysis of the influence of the location of the pollution source in the street canyon on the maximum concentration of the impurity in the breathing zone (up to 2 m from the bottom of the canyon) and in the canyon as a whole is carried out. The geometric characteristics of the street canyon are: height (H) and width (W) 20 meters, depth (L) – 30 meters. The calculations were carried out on a structured Cartesian grid 110 62 100 along the axes Ox , Oy and Oz , respectively. The mesh was compacted at the border nodes. The size of 5 border cells in the direction normal to the solid boundary was 0,1 m. This is done to get the dimensionless y^+ into the recommended range of 30 to 300.

The source of constant-intensity admixture was located in the center, at the leeward wall and at the windward wall of the street canyon, near the bottom.

The viscosity of the medium is considered to be $\nu = 15 \cdot 10^{-6} \text{ m}^2/\text{s}$, which corresponds to the viscosity of the air at 20 °C. The boundary conditions were set as follows: at the inlet boundary, the horizontal velocity $U_{in} = 1 \text{ m/s}$; at the outlet boundary, the normal velocity derivatives are equal to zero; no-slip conditions were considered on the generators of the street canyon and on the roofs of buildings; sliding conditions are set on the upper wall. Initial conditions: the longitudinal velocity is equal to the inlet velocity (1 m/s), and the vertical velocity and pressure are zero.

As can be seen from the vector velocity fields (Figure 2), the wind flow passing over the canyon affects the air masses inside the canyon, setting them in motion. Thus, one large vortex is formed within the canyon, which, in turn, acts on the source of the impurity, setting the impurity in motion. Thus, the weathering process begins.

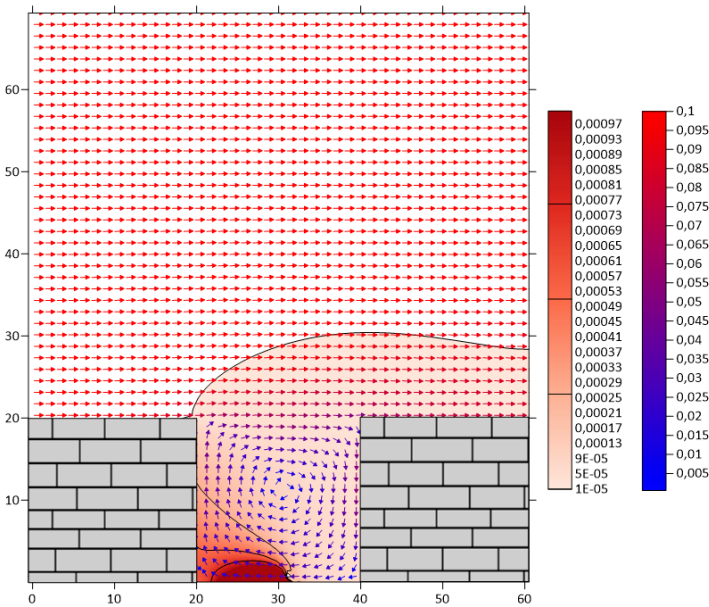


Fig. 2. Vector velocity field and the contour fields of concentration with center location of the pollution source. The width, height and depth of the street canyon are 20 m, 20 m and 30 m, respectively

As a result of modeling the transport of impurities in a street canyon, the following pictures of the distribution of its concentration were obtained (Figure 2, 3).

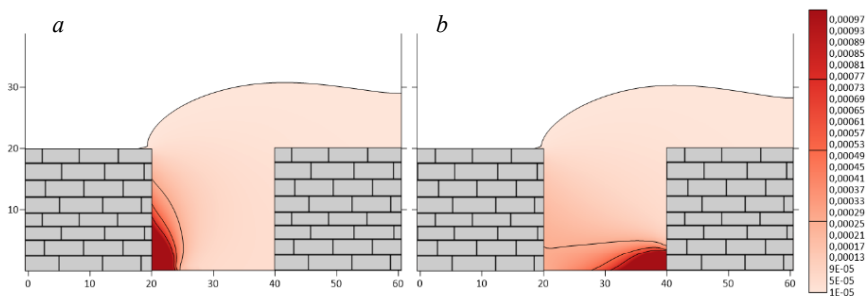


Figure 3. Contour fields of concentration in a street canyon with
a) left and *b)* right locations of the pollution source

In all three cases, the admixture is weathering from the canyon. When the source is displaced to any of the street canyon generators, the deterioration of the admixture weathering is observed. This fact is confirmed by the following integral characteristics of the maximum concentrations of impurities in the breathing zone (Table).

Maximum impurity concentrations in the breathing zone

Source location	Left	Middle	Right
<i>C</i>	1,919602	0,15369	0,51267

The maximum impurity concentrations in case of left source location are several times higher than the maximum impurity concentrations in the case of right location, which is explained by poor ventilation, which depends on the velocity and direction of the wind. In case of central source location, the impurity does not have time to accumulate and is well weathered.

Conclusion

A microscale model of turbulent air movement and impurity transport in a street canyon is presented. For isothermal turbulent air flow in the canyon, it was investigated how different variants of the location of the pollution source affect the accumulation of the incoming admixture inside the street canyon (Figure 2, 3). The integral characteristics of the impurity concentration in the

street canyon in general and in the breathing zone (up to 2 m from the bottom of the canyon) are calculated and analyzed.

The location of the source of the impurity directly depends on how the impurity will accumulate in the street canyon. The case of the central source location showed the best weathering of the admixture, the left is the worst.

REFERENCES

1. *Nuterman R., Starchenko A., Baklanov A.* Numerical model of urban aerodynamics and pollution dispersion // International Journal of Environment and Pollution. – 2011. – V. 44. – P. 385–393.
2. *Starchenko A.V., Nuterman R.B., Danilkin E.A.* Numerical study of turbulent flows and pollution transport in street canyons. – Tomsk: Tomsk University Press, 2015. – 252 p.
3. *Danilkin E.A., Starchenko A.V.* Modelling the transfer of road transport emissions in a street canyon // Computational technologies. – 2020. – V. 22(4). – P. 4–21.
4. *Henkes R.A.W.M., Van Der Flugt F. F., Hoodendoorn C. J.* Natural convection flow in a square cavity calculated with Low-Reynolds-number turbulence models // Int. J. Heat Mass Transfer. – 1991. – V. 34(2). – P. 377–388.
5. *Danilkin E.A., Lechshinskiy D.V., Starchenko A.V.* Natural Numerical research of the influence of wall heating on the turbulent flow pattern in an idealized street canyon // J. Phys.: Conf. Ser. – 2020. – 1715 012015.

Leshchinskiy Dmitriy Victorovich, junior research assistant of The Regional Scientific and Educational Mathematical Center of Tomsk State University; 360flip182@gmail.com;

Danilkin Evgeniy Alexandrovich, Candidate of Physical and Mathematical Sciences, senior research assistant The Regional Scientific and Educational Mathematical Center of Tomsk State University; ugin@math.tsu.ru;

Starchenko Alexander Vasilyevich, Doctor of Physical and Mathematical Sciences, professor, leading scientific worker of The Regional Scientific and Educational Mathematical Center of Tomsk State University; starch@math.tsu.ru.

Расчет сопряженного нестационарного теплообмена затупленных тел при их сверхзвуковом обтекании*

В.И. Зинченко, В.Д. Гольдин, В.И. Лаева

Томский государственный университет, г. Томск, Россия

Рассмотрена задача сопряженного теплообмена сферически затупленных конусов при их сверхзвуковом обтекании. Обтекание рассчитывается на основе модели трехмерного пограничного слоя. Внутри тела решается трехмерная задача теплопроводности. Показано, что использование высокотеплопроводных теплозащитных материалов может заметно снизить максимальные температуры тела вследствие перетекания тепла от лобовой части тела на боковую поверхность. На основании полученных результатов предложены обобщающие зависимости, позволяющие производить оценку снижения максимальной температуры поверхности без сложных вычислений.

Ключевые слова: *аэродинамический нагрев, сопряженный теплообмен, теплозащитные материалы, ламинарный пограничный слой.*

В последнее время развиваются новые подходы к тепловой защите сверхзвуковых летательных аппаратов, подвергающихся аэродинамическому нагреву в процессе полета. С одной стороны, это проявляется в создании новых теплозащитных материалов, способных выдерживать значительные температуры [1, 2]. С другой стороны, обозначилась тенденция использования в качестве теплозащиты материалов с высокой теплопроводностью [3]. Такие материалы обеспечивают переток тепла от лобовой части летательного аппарата на его боковую поверхность, что позволяет снизить значения максимальной температуры тела. При этом для прогнозирования теплового состояния тела требуется решение задачи сопряженного теплообмена с учетом трехмерного течения в пограничном слое и нестационарного трехмерного распространения тепла внутри теплозащитной оболочки. Целью работы является сравнение эффективности применения различных материалов для обеспечения тепловой защиты тела и заданного диапазона максимальных температур его поверхности, а также получение обобщающих зависимостей для инженерных оценок этих температур.

* Результаты были получены в рамках выполнения государственного задания Минобрнауки России, проект № 0721-2020-0032.

Постановка задачи

Рассматривается обтекание затупленного по сфере конуса с углом полураствора β , радиусом сферического затупления R_n и длиной z_k сверхзвуковым потоком газа под углом атаки α . Лобовая часть конуса выполнена из сплошного материала, а на его боковой части ($z > z_0 = 1,5$) имеется коническая оболочка постоянной толщины: $L = 0,5$ (все линейные размеры задачи отнесены к R_n). Схема расчетной области представлена на рис. 1.

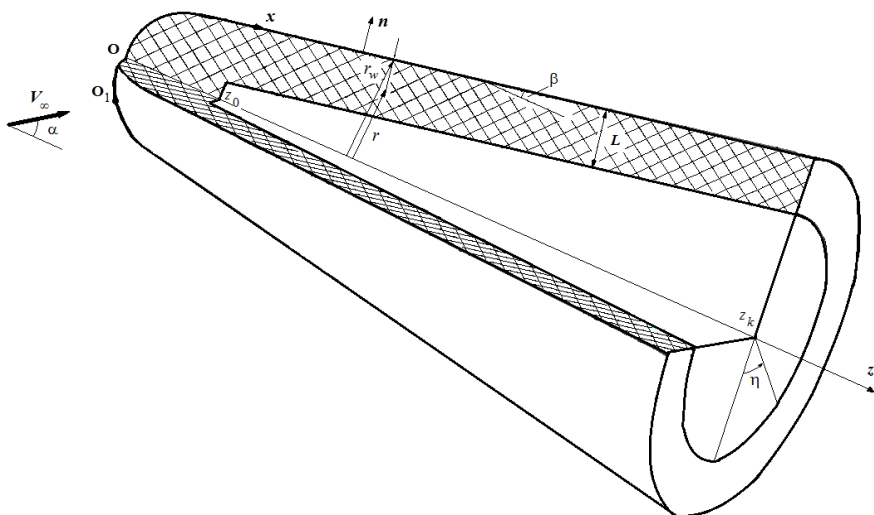


Рис. 1. Схема расчетной области

Расчет течения в трехмерном пограничном слое на поверхности конуса осуществлялся в естественной системе координат $\xi = x/R_n$, η , n , где x – длина дуги образующей поверхности тела, отсчитываемая от точки O ; η – угол, отсчитываемый от плоскости симметрии с наветренной стороны тела; n – расстояние до поверхности тела. Для определения термодинамических параметров использовалась модель равновесного состояния воздуха. Коэффициент динамической вязкости вычислялся по формуле Сазерленда. Для определения поля течения использовалась система уравнений ламинарного пограничного слоя в переменных Дородницына – Лиза [4, 5]:

$$\begin{aligned} \frac{\partial}{\partial \zeta} \left(l \frac{\partial \bar{u}}{\partial \zeta} \right) + (f + \alpha_3 \varphi) \frac{\partial \bar{u}}{\partial \zeta} &= \alpha_1 \left(\bar{u} \frac{\partial \bar{u}}{\partial \xi} - \frac{\partial f}{\partial \xi} \frac{\partial \bar{u}}{\partial \zeta} \right) + \alpha_2 \left(\bar{\omega} \frac{\partial \bar{u}}{\partial \eta} - \frac{\partial \varphi}{\partial \eta} \frac{\partial \bar{u}}{\partial \zeta} \right) + \\ &+ \beta_1 \left(\bar{u}^2 - \frac{\rho_e}{\rho} \right) + \beta_2 \left(\bar{\omega}^2 - \frac{\rho_e}{\rho} \right) + \beta_3 \left(\bar{u} \bar{\omega} - \frac{\rho_e}{\rho} \right); \\ \frac{\partial}{\partial \zeta} \left(l \frac{\partial \bar{\omega}}{\partial \zeta} \right) + (f + \alpha_3 \varphi) \frac{\partial \bar{\omega}}{\partial \zeta} &= \alpha_1 \left(\bar{u} \frac{\partial \bar{\omega}}{\partial \xi} - \frac{\partial f}{\partial \xi} \frac{\partial \bar{\omega}}{\partial \zeta} \right) + \alpha_2 \left(\bar{\omega} \frac{\partial \bar{\omega}}{\partial \eta} - \frac{\partial \varphi}{\partial \eta} \frac{\partial \bar{\omega}}{\partial \zeta} \right) + \\ &+ \beta_4 \left(\bar{\omega}^2 - \frac{\rho_e}{\rho} \right) + \beta_5 \left(\bar{u} \bar{\omega} - \frac{\rho_e}{\rho} \right); \end{aligned} \quad (1)$$

$$\begin{aligned} \frac{\partial}{\partial \zeta} \left\{ \frac{l}{\text{Pr}} \frac{\partial g}{\partial \zeta} + \gamma l \left(1 - \frac{1}{\text{Pr}} \right) \frac{1}{2} \frac{\partial}{\partial \zeta} \left[\bar{u}^2 + \left(\frac{\omega_e}{u_e} \bar{\omega} \right)^2 \right] \right\} + (f + \alpha_3 \varphi) \frac{\partial g}{\partial \zeta} &= \\ = \alpha_1 \left(\bar{u} \frac{\partial g}{\partial \xi} - \frac{\partial f}{\partial \xi} \frac{\partial g}{\partial \zeta} \right) + \alpha_2 \left(\bar{\omega} \frac{\partial g}{\partial \eta} - \frac{\partial \varphi}{\partial \eta} \frac{\partial g}{\partial \zeta} \right); \end{aligned}$$

$$p_e = \rho h \frac{\gamma_{\text{eff}} - 1}{\gamma_{\text{eff}}}.$$

Здесь

$$\begin{aligned} \zeta &= \frac{u_e r_w}{\left(2R_n \int_0^{\xi} \rho_e \mu_e u_e r_w^2 d\xi \right)^{1/2}} \int_0^{\xi} \rho dn, \quad \alpha_1 = \frac{2 \int_0^{\xi} \rho_e \mu_e u_e \bar{r}_w^2 d\xi}{\rho_e \mu_e u_e \bar{r}_w^2}, \quad \alpha_2 = \frac{\omega_e}{u_e} \frac{\alpha_1}{\bar{r}_w}, \\ \alpha_3 &= \frac{1}{\rho_e \mu_e u_e \bar{r}_w^3} \frac{\omega_e}{u_e} \int_0^{\xi} \frac{\partial}{\partial \eta} (\rho_e \mu_e u_e \bar{r}_w^2) d\xi + \frac{\alpha_1}{\bar{r}_w} \frac{\partial}{\partial \eta} \left(\frac{\omega_e}{u_e} \right), \quad \beta_1 = \frac{\alpha_1}{u_e} \frac{\partial u_e}{\partial \xi}, \\ \beta_2 &= -\frac{1}{\bar{r}_w} \frac{d\bar{r}_w}{d\xi} \alpha_1 \left(\frac{\omega_e}{u_e} \right)^2, \quad \beta_3 = \frac{\alpha_2}{u_e} \frac{\partial u_e}{\partial \eta}, \quad \beta_4 = \frac{\alpha_2}{\omega_e} \frac{\partial \omega_e}{\partial \eta}, \\ \beta_5 &= \alpha_1 \left(\frac{1}{\omega_e} \frac{\partial \omega_e}{\partial \xi} + \frac{1}{\bar{r}_w} \frac{d\bar{r}_w}{d\xi} \right), \quad \bar{u} = \frac{u}{u_e}, \quad \bar{\omega} = \frac{\omega}{\omega_e}, \quad g = \frac{H}{H_e}, \end{aligned}$$

$$f = \int_0^{\zeta} \bar{u} d\zeta, \quad \varphi = \int_0^{\zeta} \bar{\omega} d\zeta, \quad \bar{r}_w = \frac{r_w}{R_n}, \quad l = \frac{\mu\rho}{\mu_e \rho_e}, \quad \gamma = \frac{u_e^2}{2H_e},$$

r , z и η – радиальная, осевая и окружная цилиндрические координаты; u и ω – x - и η -компоненты скорости; p_e – давление газа в пограничном слое; h и H – статическая и полная энтальпии газа; ρ , μ – плотность и вязкость газа; Pr – число Прандтля ($Pr = 0,72$); γ_{eff} – эффективный показатель адиабаты газа; r_w – расстояние от точки на теле до его оси симметрии, индекс « e » относится к значениям на внешней границе пограничного слоя, « 0 » – к параметрам торможения невязкого потока.

Для расчета нестационарного теплового поля внутри теплозащитной оболочки использовалось уравнение теплопроводности, которое в безразмерной форме имеет вид (теплофизические коэффициенты считались постоянными)

$$\frac{1}{S} \frac{\partial \theta}{\partial \tau} = \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \theta}{\partial r} \right) + \frac{\partial^2 \theta}{\partial z^2} + \frac{1}{r^2} \frac{\partial^2 \theta}{\partial \eta^2}. \quad (2)$$

Здесь $\theta = T/T_{e0}$, T – температура, T_{e0} – температура торможения на внешней границе пограничного слоя; $S = \frac{\lambda_s}{\sqrt{\rho_{e0} \mu_{e0} V_m R_n} \frac{T_{e0}}{H_e}}$ – параметр

сопряженности, характеризующий отношение тепловых потоков, поступающих в тело и со стороны газовой фазы, $V_m = \sqrt{2H_e}$ и

$$\tau = \frac{\lambda_s t}{\rho_s c_s R_n^2 S}, \quad \rho_s, c_s, \lambda_s - \text{плотность, удельная теплоемкость и коэффициент теплопроводности тела.}$$

Граничные условия на внешней границе пограничного слоя имеют вид

$$\bar{u}(\xi, \eta, \infty) = \bar{\omega}(\xi, \eta, \infty) = g(\xi, \eta, \infty) = 1. \quad (3)$$

На границе раздела сред ставились граничные условия 4-го рода (включающие равенство температур и тепловых потоков в пограничном слое и в твердом теле):

$$\bar{u}(\xi, \eta, 0) = \bar{\omega}(\xi, \eta, 0) = 0, \quad f(\xi, \eta, 0) = \varphi(\xi, \eta, 0) = 0; \quad (4)$$

$$\tilde{q}_w - \pi_\sigma \theta_w^4 = -S \frac{\partial \theta}{\partial n_1}. \quad (5)$$

Здесь $\pi_{\sigma} = \frac{\varepsilon \sigma T_{e0}^4 \sqrt{R_n}}{h_{e0} \sqrt{\rho_{e0} \mu_{e0} V_m}}$, ε – излучательная способность поверхности

тела (степень черноты); $\tilde{q}_w = \frac{q_w}{q_w^*}$, $q_w^* = \sqrt{\frac{\rho_{e0} \mu_{e0} V_m}{R_n}} H_e$ — характерное значение q_w ; в переменных Дородницына – Лиза выражение для безразмерного теплового потока имеет вид:

$$\tilde{q}_w = \sqrt{\frac{u_e \rho_e \mu_e}{V_m \rho_{e0} \mu_{e0}} \frac{1}{\alpha_1} \left(\frac{l}{Pr} \frac{\partial g}{\partial \zeta} \right)_w}.$$

В качестве граничных условий для уравнения (2) в плоскостях симметрии $\eta = 0$, π задавалось условие симметрии для температуры, а на внутренней поверхности оболочки тела и в ее тыльной части – условие тепловой изоляции

$$\frac{\partial \theta}{\partial n} = 0, \quad (6)$$

где дифференцирование ведется по нормали к соответствующей поверхности.

В начальный момент времени температура тела предполагалась постоянной:

$$\theta(0, r, z, \eta) = \theta_{in} = \frac{T_{in}}{T_{e0}}. \quad (7)$$

Решение задачи теплопроводности (2) зависит от параметра сопряженности S и параметра π_{σ} , характеризующего влияние излучения поверхности тела. При $S = 0$ нет необходимости решать уравнение теплопроводности, а решение уравнений пограничного слоя с граничным условием (5) (при $S = 0$) дает распределение радиационно-равновесной температуры по поверхности тела. Случай $S \rightarrow \infty$ соответствует материалу с бесконечной теплопроводностью. При этом температура тела будет зависеть только от времени. Уравнение для ее определения можно получить путем интегрирования (2) по объему оболочки с учетом граничных условий (5), (6) аналогично тому, как это делается в работе [6]:

$$\Omega \frac{d\theta_w}{d\tau} = \int_{\Sigma} \tilde{q}_w d\Sigma - \pi_{\sigma} \Sigma \theta_w^4, \quad (8)$$

где Ω – безразмерный объем оболочки, отнесенный к R_n^3 ; Σ – безразмерная площадь поверхности раздела сред, отнесенная к R_n^2 . При решении уравнения (8) необходимо иметь в виду, что значение \tilde{q}_w зависит от τ и его нужно определять на каждом шаге по времени.

Методика решения

Распределение давления вдоль по поверхности тела определялось из невязкой задачи обтекания с использованием методов С.К. Годунова и Мак-Кормака [7]. Поле течения в пограничном слое рассчитывалось методом И.В. Петухова [8]. Уравнение теплопроводности (2) решалось методом «Ромб» Гаджиева – Писарева [9], обобщенным на трехмерный случай [10].

Результаты расчета тепловых потоков при заданной температуре поверхности сравнивались с результатами экспериментов в аэродинамических трубах [11] при следующих параметрах: $R_n = 0,0279$ м, $\beta = 15^\circ$, число Маха в набегающем потоке $M_\infty = 10,6$, параметры торможения $T_{e0} = 1111$ К и $p_{e0} = 0,171$ бар, температура поверхности $T_w = 294$ К, угол атаки менялся от 0 до 10° . Сравнение показало удовлетворительное согласие расчетов и экспериментов (рис. 2)

Алгоритм решения сопряженной задачи состоял из следующих этапов:

1. При заданных условиях набегающего потока из решения задачи невязкого обтекания рассчитывалось распределение давления по поверхности тела и параметры на внешней границе пограничного слоя.

2. Для заданного распределения температуры поверхности тела из решения уравнений пограничного слоя определялись тепловой поток к телу и коэффициент теплообмена

$$\left(\frac{\alpha}{c_p}\right) = \frac{q_w(\xi, \eta)}{H_e - h_w(\xi, \eta)}.$$

3. При вычисленном на предыдущем этапе коэффициенте теплообмена решалась задача теплопроводности внутри тела с граничными условиями 3-го рода; в результате определялось распределение температуры на следующем шаге по времени.

4. Этапы 3, 4 повторялись до достижения стационарных значений температуры тела.

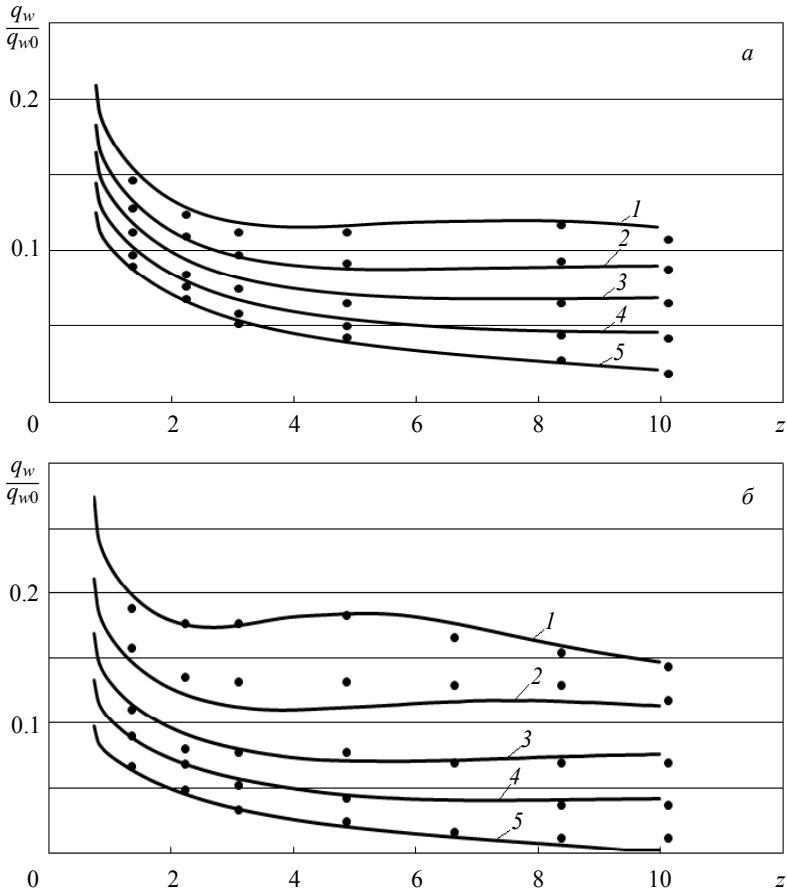


Рис. 2. Распределение теплового потока на боковой поверхности затупленного конуса ($\beta = 15^\circ$), обтекаемого под углами атаки $\alpha = 5^\circ$ (а) и 10° (б), при $\eta = 0^\circ$ (кр. 1), 60° (кр. 2), 90° (кр. 3), 120° (кр. 4) и 180° (кр. 5): точки – экспериментальные данные [11]

Осесимметричное течение, $\alpha = 0^\circ$

Для случая обтекания тела под нулевым углом атаки была проведена серия расчетов сопряженной задачи при следующих параметрах: $M_\infty = 6,1$, $p_{e0} = 2,2$ бар, $T_{in} = 293$ К, $\varepsilon = 0,8$, $\beta = 5^\circ$, $L = 0,5$, $z_k = 20$, $z_0 = 1,5$. Исследовалось влияние геометрических факторов $R_n = 0,004$, $0,01$ и $0,02$ м, а также температур торможения $T_{e0} = 562$, 1051 , 1549 , 1948

и 2538 К. В расчетах в качестве тепловой защиты использовались материалы с теплофизическими характеристиками, приведенными в таблице. Все вычисления проводились до достижения стационарных температурных полей внутри тела.

Материал	$\lambda_s, \text{Вт}/(\text{м} \cdot \text{К})$	$\rho_s, \text{кг}/\text{м}^3$	$c_s, \text{Дж}/(\text{кг} \cdot \text{К})$
Сталь	20	7800	600
Алюминий-1 (Al-1)	170	2800	1000
Алюминий-2 (Al-2)	240	2300	1000
Медь	386	8950	370

На рис. 3, 4 приведены результаты решения задачи при $R_n = 0,004 \text{ м}$, $T_{e0} = 562 \text{ К}$.

Распределения температуры поверхности для разных материалов в момент времени $t = 20$ и 500 с показаны на рис. 3; там же показаны значения температуры для материала с бесконечной теплопроводностью ($S \rightarrow \infty$, кривые 4, 8) и нулевой теплопроводностью ($S = 0$, радиационно-

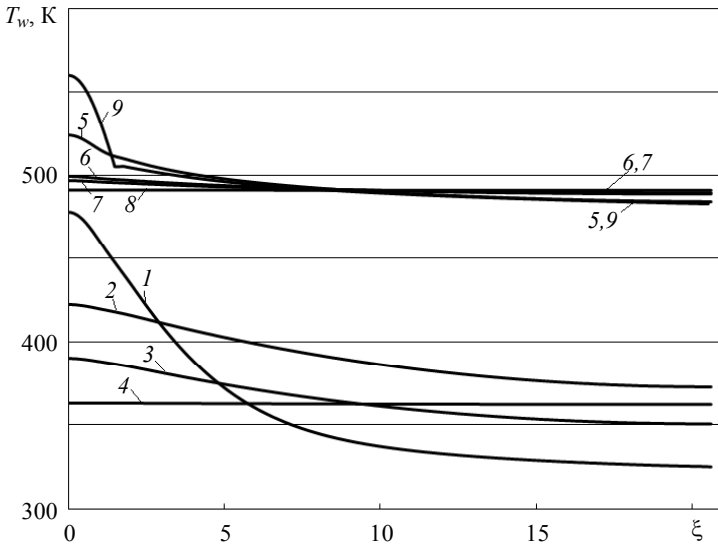


Рис. 3. Температура поверхности тела в различные моменты времени: кр. 1–4 соответствуют $t = 20 \text{ с}$, кр. 5–8 относятся к $t = 500 \text{ с}$, кр. 9 – радиационно-равновесная температура; кр. 1, 5 – сталь, кр. 2, 6 – Al-2, кр. 3, 7 – медь, кр. 4, 8 – материал с бесконечной теплопроводностью

равновесная температура, кривая 9). Для всех материалов максимальная температура достигается в передней точке тела, однако ее значение меняется в зависимости от используемого теплозащитного материала. Увеличение теплопроводности материала выравнивает распределение стационарной температуры по поверхности тела, уменьшает температуру на лобовой поверхности, где достигается ее максимум, и увеличивает на боковой. Диапазон изменения максимальной температуры определяется двумя указанными выше предельными случаями. В рассмотренном случае выбор высокотеплопроводного материала может снизить ее значение на 12%.

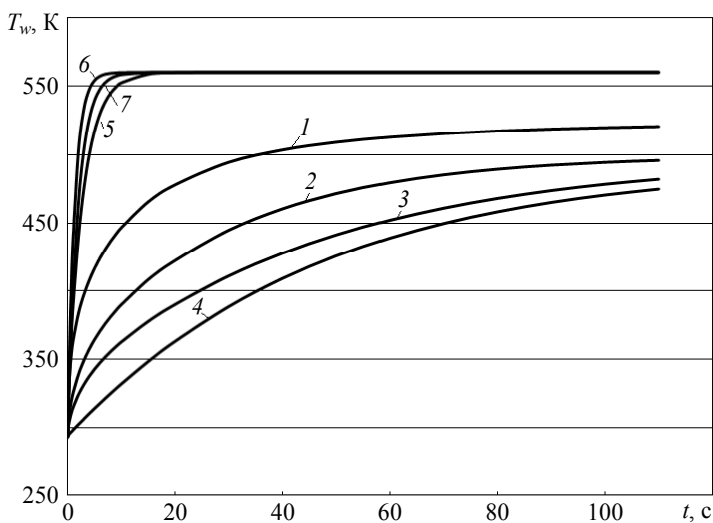


Рис. 4. Температура в критической точке в зависимости от времени: кр. 1–4 относятся к двумерному расчету, кр. 5–7 – к одномерному; кр. 1, 5 – сталь, кр. 2, 6 – Al-2, кр. 3, 7 – медь, кр. 4 – материал с бесконечной теплопроводностью

На рис. 4 приведены значения максимальной температуры поверхности в зависимости от времени. Увеличение теплопроводности, с одной стороны, снижает эту температуру и на нестационарном участке прогрета, а с другой, увеличивает время выхода на стационарные значения. Кривыми 5 – 7 показаны значения температуры, полученные из решения одномерной задачи теплопроводности, когда в правой части уравнения (2) при $r \rightarrow 0$ оставляется только вторая производная по z .

Одномерное решение гораздо раньше достигает стационарного радиационно-равновесного значения. Учет перетекания тепла на боковую поверхность на нестационарном участке сказывается гораздо сильнее и может снижать максимальную температуру на 40% по сравнению с одномерным случаем.

Как показывают расчеты, проведенные вплоть до значений температуры торможения $T_{e0} = 2538$ К, с ростом температуры влияние переизлучения поверхности (параметра π_σ) возрастает, и возрастает эффективность использования высокотеплопроводных материалов для снижения максимальных температур поверхности. Возможные пределы такого снижения можно оценить из рис. 5, где показано значение безразмерной максимальной температуры поверхности $\theta_{w,st}$ в зависимости от параметра π_σ для двух предельных случаев тепловой защиты: теплоизоляционного материала ($S = 0$, θ_{wr}) и материала с бесконечной теплопроводностью ($S \rightarrow \infty$). Значение температуры для реального материала ($S = 6$) лежит между кривыми 1 и 2. На рис. 5 кривой 4 показана разность между предельными значениями температуры.

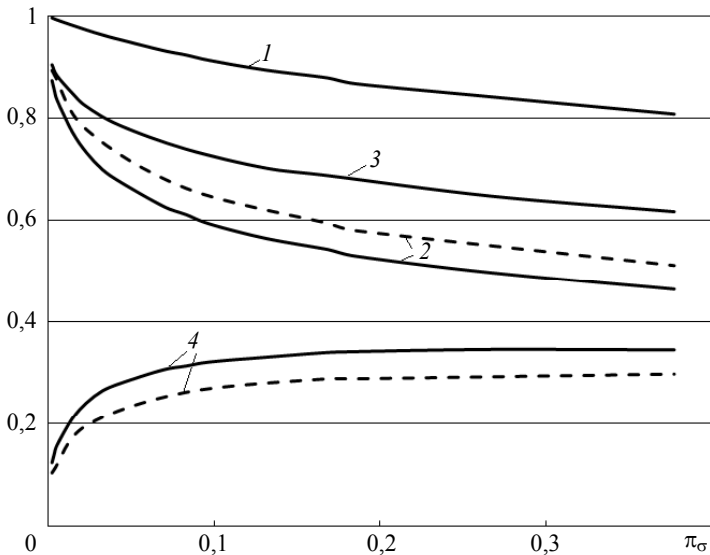


Рис. 5. Значения стационарной температуры в критической точке. Кр. 1 соответствует θ_{wr} ($S = 0$), кр. 2 – θ_w ($S \rightarrow \infty$), кр. 3 – θ_w ($S = 6$), кр. 4 – $\theta_{wr} - \theta_w$ ($S \rightarrow \infty$), пунктиром показаны значения при $z_k = 10$

Поскольку значение максимальной стационарной температуры поверхности лежит в указанных пределах, то для обобщения полученных результатов удобно ввести величину ($0 < \varphi_{st} < 1$)

$$\varphi_{st} = \frac{\theta_{wr} - \theta_w}{\theta_{wr} - \theta_w (S \rightarrow \infty)}, \tag{9}$$

где θ_w соответствует реальному материалу, θ_{wr} – радиационно-равновесной температуре, $\theta_w (S \rightarrow \infty)$ – материалу с бесконечной теплопроводностью.

На рис. 6, 7 приведены зависимости этой величины от основных параметров задачи S и π_σ .

Кривые 1 – 4 на рис. 6 получены при различных значениях π_σ , которые определялись по следующим значениям T_{e0} , R_n : кр. 1 – $\pi_\sigma = 0,246 \cdot 10^{-2}$, $T_{e0} = 562$ К, $R_n = 0,004$ м; кр. 2 – $\pi_\sigma = 0,0233$, $T_{e0} = 1051$ К, $R_n = 0,01$ м; кр. 3 – $\pi_\sigma = 0,131$, $T_{e0} = 1948$ К, $R_n = 0,01$ м; кр. 4 – $\pi_\sigma = 0,377$, $T_{e0} = 2538$ К, $R_n = 0,02$ м. Как видно из рисунка, кривые начинают сближаться при росте π_σ (росте температуры торможения).

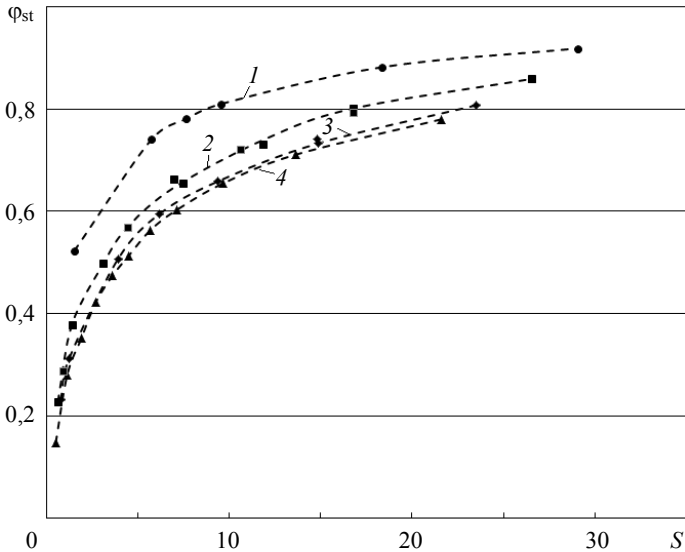


Рис. 6. Критическая точка, стационарные значения температуры. S – параметр сопряженности: ● – $T_{e0} = 562$ К, ■ – $T_{e0} = 1051$ К, ◆ – $T_{e0} = 1948$ К, ▲ – $T_{e0} = 2538$ К

Этот эффект подтверждается также рис. 7, где показаны зависимости φ_{st} от π_σ при различных S – при возрастании π_σ эти зависимости близки к слабо меняющимся линейным.

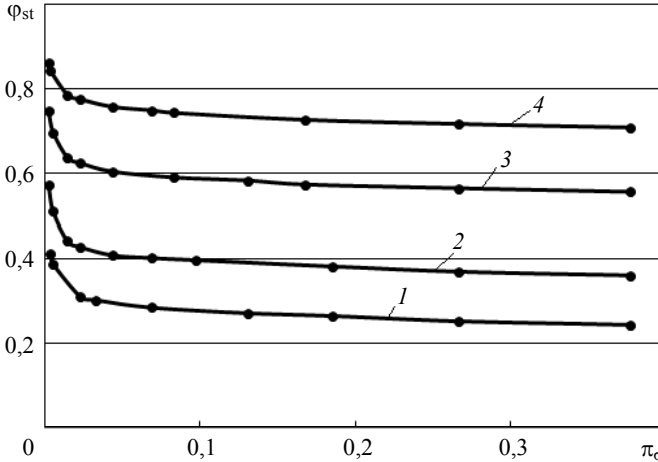


Рис. 7. Величина φ_{st} в зависимости от π_σ при различных значениях S . Кр. 1 соответствует $S = 1$, кр. 2 – $S = 2$, кр. 3 – $S = 6$, кр. 4 – $S = 15$

Используя результаты, приведенные на рис. 6 можно предложить следующий алгоритм приближенного вычисления максимальной стационарной температуры: для заданного теплозащитного материала определяется значение параметра сопряженности S ; затем по рис. 6 и опорному значению π_σ находится значение φ_{st} ; если реальное π_σ отличается от опорного, то можно использовать линейную интерполяцию; затем из (9) находится безразмерное значение температуры при известной радиационно-равновесной температуре и температуре материала с бесконечной теплопроводностью, которая вычисляется путем решения уравнения (8); по безразмерной температуре можно легко найти абсолютную температуру.

Можно получить аналогичную обобщенную зависимость и для нестационарного значения температуры поверхности в передней критической точке. Для этой цели нужно использовать следующую безразмерную величину:

$$\varphi_n = \frac{\theta_{w,1}(\tau, S) - \theta_{w,2}(\tau, S)}{\theta_{w,1}(\tau, S) - \theta_w(\tau, S \rightarrow \infty)} \approx \frac{\theta_{wr} - \theta_{w,2}(\tau, S)}{\theta_{wr} - \theta_w(\tau, S \rightarrow \infty)}, \quad (10)$$

где $\theta_{w,2}(\tau, S)$, $\theta_{w,1}(\tau, S)$ – значения температуры, полученные в двумерном и одномерном расчете, $\theta_w(\tau, S \rightarrow \infty)$ – значение, полученное из решения уравнения (8). Приближенное равенство в (10) имеет малую погрешность после небольшого начального промежутка времени ($\tau > 2$). Анализ поведения φ_n в зависимости от времени показывает, что эта величина быстро выходит на свое стационарное значение и поэтому ее можно заменить на φ_{st} , которая вычисляется по формуле (9). В результате приближенное значение максимальной температуры поверхности в зависимости от времени будет определяться следующим образом:

$$\theta_{w,2}(\tau, S) = \theta_{w,1}(\tau, S) - \varphi_{st} [\theta_{w,1}(\tau, S) - \theta_{w,2}(\tau, S \rightarrow \infty)]. \quad (11)$$

На рис. 8 приведены сравнения значений температуры $\theta_w(\tau, S)$, полученных из численного решения двумерной задачи (сплошные кривые) и по приближенной методике, использующей формулу (11) (пунктирные линии).

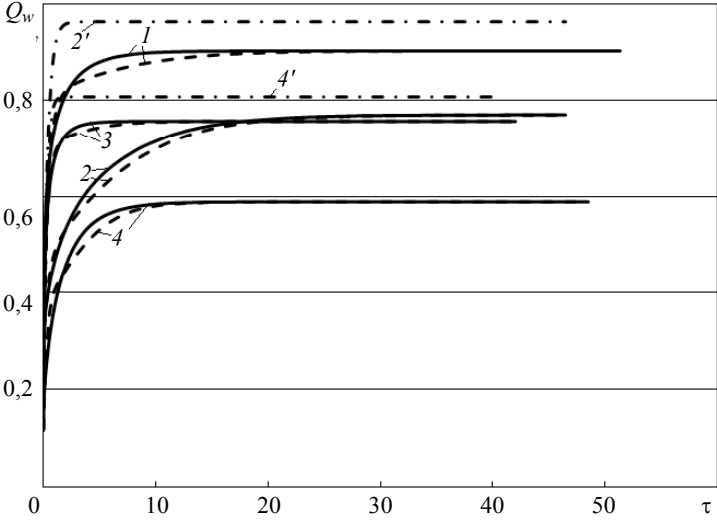


Рис. 8. Температура поверхности в критической точке в зависимости от времени: кр. 1 – $\pi_\sigma = 0,033$, $S = 0,62$; кр. 2 – $\pi_\sigma = 0,033$, $S = 11,9$; кр. 3 – $\pi_\sigma = 0,377$, $S = 0,51$; кр. 4 – $\pi_\sigma = 0,377$, $S = 9,7$. Сплошные кривые 1–4 соответствуют двумерному расчету, пунктирные – приближенной методике; кр. 2' – одномерный расчет при $\pi_\sigma = 0,033$, $S = 11,9$; кр. 4' – одномерный расчет при $\pi_\sigma = 0,377$, $S = 9,7$

Влияние угла атаки (трехмерный расчет)

Для изучения влияния угла атаки были проведены расчеты при следующих исходных данных: $M_\infty = 6,1$, $p_{e0} = 2,2$ бар, $T_{in} = 293$ К, $\varepsilon = 0,8$, $L = 0,5$, $z_k = 10$, $z_0 = 1,5$, $\beta = 10^\circ$. Расчеты проводились для радиусов затупления $R_n = 0,004, 0,01, 0,02$ м, температуры торможения $T_{e0} = 500, 1000, 1500, 2000$ К и углов атаки $\alpha = 5, 10^\circ$. Использовались теплофизические характеристики материалов поверхности, приведенные в таблице. Так же, как и в двумерном случае, эволюция трехмерного температурного поля тела в каждом расчете прослеживалась до достижения стационарного состояния.

На рис. 9 показаны распределения температуры по поверхности тела в плоскости симметрии при $T_{e0} = 1500$ К, $R_n = 0,004$ м и угле атаки $\alpha = 10^\circ$ для различных материалов поверхности. Как и следовало ожидать, увеличение угла атаки до 10° приводит к росту T_w на наветренной стороне периферийной части конуса примерно на 30–100 К в стационарном режиме по сравнению с осесимметричным обтеканием. На подветренной стороне наблюдается снижение T_w (на 30 К для стали). В то же время максимальная

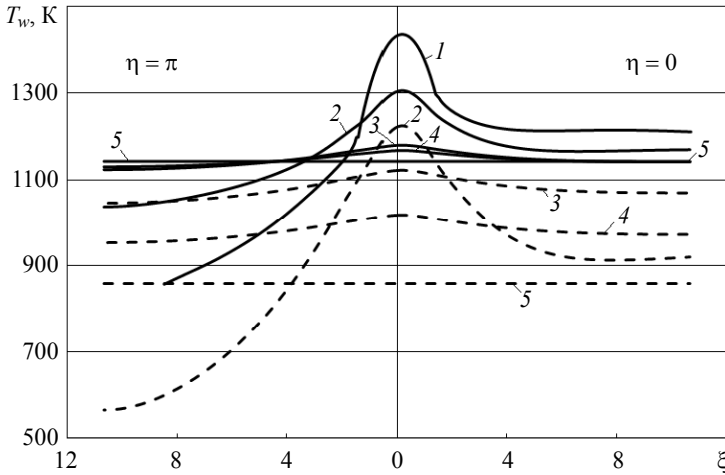


Рис. 9. Распределение температуры по поверхности тела вдоль плоскости симметрии при $\alpha = 10^\circ$: кр. 1 – радиационно-равновесная температура ($\lambda_s = 0$); кр. 2 – сталь; кр. 3 – AL-2; кр. 4 – медь; кр. 5 – материал с бесконечной теплопроводностью ($\lambda_s \rightarrow \infty$); сплошные линии – стационарные значения параметров; штриховые линии – значения температуры в момент времени $t = 20$ с

температура, которая достигается в окрестности критической точки, практически не меняется для менее теплопроводной стали и увеличивается примерно на 30 К в предельном случае при $\lambda_s \rightarrow \infty$. Из рис. 9 видно, что максимальная температура поверхности тела может быть снижена (до 25%) за счет выбора материала тепловой защиты.

Для представления пространственной картины распределения температур по окружной координате на поверхности тела на рис. 10 показаны зависимости безразмерной температуры θ_w от переменной η при двух значениях ξ при выходе на стационарный режим. Как видно из рис. 9, 10 температура наветренной и подветренной сторон заметно различается только для низко теплопроводной стали.

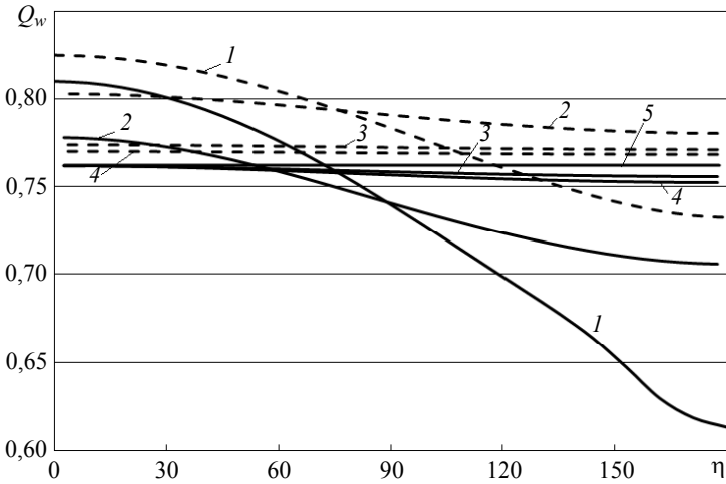


Рис. 10. Распределения безразмерной температуры поверхности тела по окружной координате при $\alpha = 10^\circ$, $\xi = 2$ (штриховые линии) и 7 (сплошные линии); обозначения линий такие же, как на рис. 9

Поскольку максимальная температура поверхности достигается в окрестности критической точки течения и она мало зависит от угла атаки, то для приближенного вычисления ее стационарного значения можно использовать методику, основанную на применении формулы (9) для осесимметричного течения. Кроме того, расчеты показывают, что и при наличии угла атаки для определения зависимости от времени максимальной температуры поверхности можно использовать формулу (11).

Заключение

Таким образом, путем проведения большого количества расчетов сопряженного нестационарного теплообмена тела при его сверхзвуковом обтекании в широком диапазоне определяющих параметров оценена возможность влияния на уровень максимальных температур тела путем выбора материалов тепловой защиты. Показано, что при использовании высокотеплопроводных материалов снижение температуры может достигать 40% на нестационарном участке прогрева и 25% – в стационарном случае при ламинарном режиме течения в пограничном слое. Наличие угла атаки в рассмотренном диапазоне слабо влияет на эти характеристики. Предложена простая методика приближенного определения максимальной температуры без проведения трудоемких двумерных и трехмерных расчетов.

СПИСОК ЛИТЕРАТУРЫ

1. *Cecere A., Savino R., Allonis C., Monteverde F.* Heat Transfer in Ultra-High-Temperature Advanced Ceramics under High Enthalpy Arc-jet Conditions // *Int. Journal of Heat and Mass Transfer.* – 2015. – V. 91. – P. 747–755.
2. *Резник С.В., Колесников А.Ф., Просунцов П.В., Гордеев А.Н., Михайловский К.В.* Обработка элементов многоразового теплозащитного покрытия из углерод-керамического композиционного материала. 2. Тепловые испытания образцов материала // *Инженерно-физический журнал.* – 2019. – Т. 92. – № 2. – С. 322–329.
3. *Гешеле В.Д., Полежаев Ю.В., Раскатов И.П., Стоник О.Г., Габбасова Г.В.* Возможности повышения скорости полета гиперзвуковых летательных аппаратов // *Теплофизика высоких температур.* – 2013. – Т. 51. – № 5. – С. 798–800.
4. *Зинченко В.И.* Математическое моделирование сопряженных задач теплообмена. – Томск: Изд-во Том. ун-та, 1985.
5. *Зинченко В.И., Фёдорова О.П.* Численное исследование пространственного ламинарного пограничного слоя с учетом сопряженного теплообмена // *Прикладная механика и техническая физика.* – 1988. – Т. 29. – № 2. – С. 34–42.
6. *Зинченко В.И., Лаева В.И., Сандрыкина Т.С.* Расчет температурных режимов обтекаемых тел с различными теплофизическими характеристиками // *Прикладная механика и техническая физика.* – 1996. – Т. 37. – № 5. – С. 105–114.
7. *Антонов В.А., Гольдин В.Д., Пахомов Ф.М.* Аэродинамика тел со вдувом. – Томск: Изд-во Том. ун-та, 1990.
8. *Петухов И.В.* Численный расчет двумерных течений в пограничном слое // *Численные методы решения дифференциальных и интегральных уравнений и квадратурные формулы.* – М.: Наука, 1964. – С. 304–324.
9. *Гаджиев А.Д., Писарев В.Н., Шестаков А.А.* Метод расчета двумерных задач теплопроводности на неортогональных сетках // *Журнал вычислительной математики и математической физики.* – 1982. – Т. 22. – № 2. – С. 339–347.

10. *Зинченко В.И., Гольдин В.Д., Зверев В.Г.* Численное исследование тепловой защиты гиперзвуковых летательных аппаратов // Теплофизика и аэромеханика. – 2018. – Т. 25. – № 3 – С. 373–381.
11. *Cleary J.W.* Effects of Angle of Attack and Bluntness on Laminar Heating-Rate Distributions of a 15° Cone at a Mach Number of 10.6 // NASA. – 1969. – TN D-5450.

Зинченко Владислав Иванович, докт. физ.-мат. наук, профессор; vdg@math.tsu.ru;
Гольдин Виктор Данилович, ст. науч. сотр. НИИ ПММ ТГУ; vdg@math.tsu.ru;
Ласва Валентина Ивановна, ст. преподаватель кафедры вычислительной математики и компьютерного моделирования ММФ НИ ТГУ; vdg@math.tsu.ru.

Математические модели и численные методы для расчета взаимодействия двух осевых вентиляторов

Е.И. Гурина

Томский государственный университет, г. Томск, Россия

При проектировании систем вентиляции возникает необходимость установки ряда осевых вентиляторов для обеспечения необходимого расхода воздуха при работе на данную сеть. Проведение многовариантных газодинамических исследований различных вариантов конструкций помогает решить задачу обоснованного выбора таких установок. Рассматривается численное моделирование взаимодействия двух осевых вентиляторов. Газодинамическое исследование течений в 3D-модели участка вентиляционной сети проведено с помощью универсального программного комплекса Ansys Fluent на основе метода контрольного объема.

Ключевые слова: моделирование турбулентности, декомпозиция расчетной области, уравнения Навье – Стокса, RANS, Ansys Fluent, газодинамическое исследование.

Течения жидкостей и газов, наблюдаемые в природных условиях и технических устройствах, делятся на ламинарные (спокойные, плавные) и турбулентные (бурные, беспорядочные) [1]. Турбулентное движение несжимаемой жидкости описывается системой уравнений, содержащей уравнения Навье – Стокса и уравнение неразрывности:

$$\frac{\partial u_i}{\partial x_i} = 0, \quad i = 1, 2, 3 ; \quad (1)$$

$$\frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j} = -\frac{1}{\rho} \frac{\partial p}{\partial x_i} + \frac{\partial}{\partial x_j} \left(\nu \frac{\partial u_i}{\partial x_j} \right), \quad i = 1, 2, 3, \quad (2)$$

где ρ – плотность; t – время; u_i – компоненты мгновенного поля скорости; ν – коэффициент кинематической вязкости; p – мгновенное значение давления; j – индекс суммирования.

В настоящее время существует три основных подхода к моделированию турбулентных течений. Первый подход (Direct Numerical Simulation, DNS), или прямое численное моделирование, связан с непо-

средственным решением вышеуказанной системы дифференциальных уравнений, аппроксимированных конечными разностями высокого порядка. Данный метод предполагает решение дискретного аналога уравнений системы без привлечения дополнительных уравнений (турбулентных замыканий). Сетка в этом случае должна быть достаточно мелкой для расчета мельчайших турбулентных вихрей, а также достаточно малым должен быть шаг по времени для разрешения высокочастотных турбулентных пульсаций. Прямое численное моделирование хорошо подходит для исследования течений при сравнительно небольших числах Рейнольдса (течения в каналах, пограничных слоях). Так как все мельчайшие турбулентные структуры (весь спектр масштабов турбулентности Колмогорова) должны быть разрешены [2], метод прямого численного моделирования является весьма ресурсоемким. В силу своей вычислительной сложности DNS почти не имеет практического применения для решения инженерных задач, в особенности промышленных.

Второй подход – моделирование крупных вихрей (Large Eddy Simulation, LES). Основная идея метода LES заключается в формальном математическом разделении крупных и мелких турбулентных структур посредством использования системы так называемых фильтрующих уравнений. Большие вихри в LES модели рассчитываются явно, а маленькие моделируются. В отличие от метода DNS, метод моделирования крупных вихрей предъявляет меньшие требования к разрешению расчетной сетки. Несмотря на современное увеличение вычислительных мощностей, использование вихреразрешающего моделирования при решении практических инженерных задач крайне ограничено.

Третий подход моделирования турбулентности требует значительно меньше вычислительных ресурсов по сравнению с методами DNS и LES. Здесь для описания развитого турбулентного течения предложена методология, согласно которой переменные в уравнениях Навье – Стокса (1), (2) делятся на осредненные и пульсационные составляющие. Случайные характеристики потока (скорость, давление, плотность) заменяются суммами осредненных и пульсационных составляющих. Например, для компонент скорости:

$$u_i(x, t) = \bar{u}_i(x, t) + u_i'(x, t), \quad i = 1, 2, 3,$$

где \bar{u}_i и u_i' – осредненная во времени и пульсационная составляющие скорости соответственно.

В рамках данного подхода течение рассчитывается с использованием осредненных по Рейнольдсу (выведены О. Рейнольдсом в 1895 г.) уравнений Навье – Стокса (Reynolds-averaged Navier-Stokes, RANS) [3]:

$$\frac{\partial \bar{u}_i}{\partial t} + \frac{\partial \bar{u}_i \bar{u}_j}{\partial x_j} = -\frac{1}{\rho} \frac{\partial \bar{p}}{\partial x_i} + \frac{\partial}{\partial x_j} \left(\nu \frac{\partial \bar{u}_i}{\partial x_j} \right) - \frac{\partial}{\partial x_j} (\overline{u_i' u_j'}), \quad (3)$$

$$\frac{\partial \bar{u}_j}{\partial x_j} = 0, \quad (4)$$

где \bar{u}_i и \bar{p} – осредненные по времени значения скорости и давления соответственно; u_i' – пульсация скорости; $\overline{u_i' u_j'}$ – напряжения Рейнольдса; ν – кинематическая вязкость; ρ – плотность среды.

Особенностью осредненных во времени уравнений Навье – Стокса (3), (4) является то, что в них содержатся нелинейные члены, включающие напряжения Рейнольдса, учитывающие дополнительные потери и перераспределение энергии в турбулентном потоке. Система уравнений Рейнольдса является незамкнутой и решить так называемую «проблему замыкания» (определение напряжений Рейнольдса) можно используя различные модели турбулентности, которые аппроксимируют $\overline{u_i' u_j'}$. Модели турбулентности могут быть алгебраическими или иметь вид дифференциальных уравнений для определения необходимых характеристик и по своей сути являются математическими моделями, устанавливающими связь между турбулентными напряжениями Рейнольдса и осредненными характеристиками течения.

Стремительный рост вычислительных мощностей оказывает существенное влияние на все сферы современной науки и техники. Возникают новые задачи, связанные с расчетом турбулентных гидродинамических течений [4], инженерные расчеты которых традиционно сопровождаются моделями, использующими осредненные характеристики (RANS-модели).

Постановка задачи

Объектом исследования в данной работе является осевой вентилятор при его работе с сетью (рис. 1). В работе [5] проведено математическое моделирование процессов, протекающих в проточной части данного осевого вентилятора встречного вращения, с помощью программного комплекса Ansys Fluent. Расчет проводился по известному заранее (данные эксперимента по статическому давлению) граничному условию на выходе из вентилятора P_{st} . В качестве расчетных параметров выступали объемный расход воздуха Q_v и значение полного давления P_{tot} на выходе из лопаточной машины. Например, для одного из режимов работы вентилятора по результатам численного моделирования были получены следующие значения аэродинамических характеристик: $Q_v = 11.37 \text{ м}^3/\text{с}$, $P_{tot} = 4403.42 \text{ Па}$ при известном избыточном статическом давлении $P_{st} = 2562 \text{ Па}$. При работе с сетью для определения параметров вентилятора необходимо знать сопротивление сети, которое определяет необходимый перепад давления, развиваемый вентилятором [6]. В случае, если на выходе из вентилятора непосредственно срез атмосферы (динамическая составляющая давления не используется за вентилятором), то в качестве перепада необходимо использовать разницу статического давления на выходе и полного давления на входе. Если же за вентилятором располагаются цилиндрические патрубки или воздуховод с плавным торможением скорости (используется динамический напор, развиваемый вентилятором), то в качестве расчетного перепада следует принимать перепад полного давления.

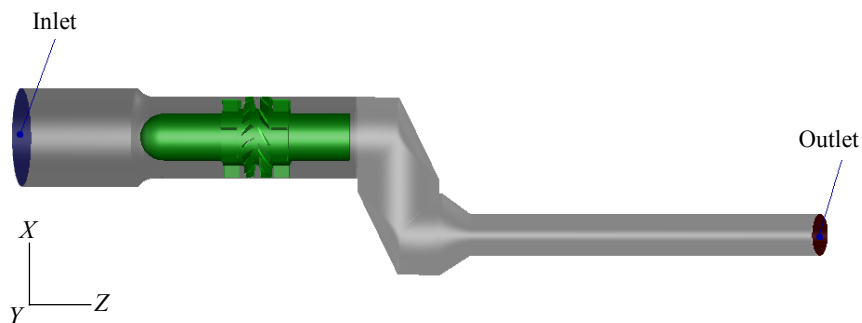


Рис. 1. Геометрия участка вентиляционной сети

В рассматриваемой установке за вентилятором воздушный поток при движении подвергается внезапному сужению после двух поворотов под прямыми углами. В рамках рассматриваемой сети поток воздуха за вентилятором имеет существенную неравномерность по радиусу и остаточную закрутку (окружная составляющая скорости). Суммарное сопротивление подобного участка сети в ряде случаев не может быть просто определено суммой сопротивлений отдельных участков [7]. Взаимная работа вентилятора и сети может быть смоделирована в ходе газодинамического анализа вентилятора с сетью.

Физико-математическая постановка

Протекающий газодинамический процесс при совместной работе вентилятора и сети рассматривается как процесс трехмерного стационарного турбулентного изотермического течения несжимаемой среды (воздуха) с заданным значением массового расхода воздуха на выходе (Outlet) и значением полного давления, равного атмосферному, на входе (Inlet). Задача – определить перепад полного давления при работе вентилятора совместно с сетью.

Забор воздушных масс идет с атмосферы и направляется далее в участок вентиляционной сети. Проведенные расчеты показали, что задание атмосферного давления сразу на входе в модель является некорректным. Небольшие изменения статического давления на входе вызывают существенное перестроение потока. Для корректного задания ГУ на входе была добавлена дополнительная расчетная зона (патрубок равного с кожухом вентилятора диаметра, длиной 1000 мм).

Математическая модель, описывающая течение потока воздуха, в данном случае включает в себя осредненные по Рейнольдсу уравнение неразрывности и уравнения Навье – Стокса (течение стационарное). Для замыкания системы уравнений (3), (4) используется градиентно-диффузионная гипотеза Буссинеска совместно с двухпараметрической стандартной $k - \varepsilon$ -моделью турбулентности. Транспортные уравнения для определения кинетической энергии турбулентности k и скорости ее диссипации ε имеют вид

$$\frac{\partial \bar{u}_j k}{\partial x_j} = \frac{\partial}{\partial x_j} \left(\left(v + \frac{v_\tau}{\sigma_k} \right) \frac{\partial k}{\partial x_j} \right) + G_k - \varepsilon, \quad v_\tau = C_\mu \frac{k^2}{\varepsilon}, \quad (5)$$

$$\frac{\partial \bar{u}_j \varepsilon}{\partial x_j} = \frac{\partial}{\partial x_j} \left(\left(\nu + \frac{\nu_\tau}{\sigma_\varepsilon} \right) \frac{\partial \varepsilon}{\partial x_j} \right) + \frac{\varepsilon}{k} C_{\varepsilon 1} G_k - C_{\varepsilon 2} \frac{\varepsilon^2}{k}; \quad (6)$$

$$-\overline{u_i' u_j'} = \nu_\tau \left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right) - \frac{2}{3} k \delta_{ij}, \quad G_k = \nu_\tau \left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right) \frac{\partial \bar{u}_i}{\partial x_j}, \quad (7)$$

где $C_{\varepsilon 1} = 1,44$; $C_{\varepsilon 2} = 1,92$; $C_\mu = 0,09$; $\sigma_k = 1,0$; $\sigma_\varepsilon = 1,3$ – коэффициенты модели турбулентности, полученные опытным путем для течения в области пограничного слоя [2]; G_k – производство турбулентной кинетической энергии; ν_τ – турбулентная вязкость; δ_{ij} – символ Кронекера.

В качестве начального приближения в расчетном домене принимается однородное распределение значений следующих характеристик: $u_i = 0$, $P_0 = 101325$ Па. Граничные условия для рассматриваемой области принимаются следующими:

- массовый расход воздуха на выходе $Q_{v_outlet} = 11,37$ м³/с;
- атмосферное давление на входе $P_{tot_inlet} = 101325$ Па;
- условие прилипания на стенках $u = 0$.

Получение разностной сетки

Численное решение системы дифференциальных уравнений в частных производных осуществляется на основе метода конечного объема с использованием универсального газодинамического программного комплекса Ansys Fluent [8]. Для согласования полей скорости и давления на начальном этапе использовался метод SIMPLE, далее Coupled. Схема аппроксимации по пространству второго порядка точности.

При проведении газодинамических расчетов лопаточных машин следует придерживаться определенной концепции по отношению к расчетной конечно-объемной модели. Должна иметься возможность быстрой замены любого под-домена с целью проведения геометрической оптимизации какого-либо участка. Следует отдать предпочтение структурированным (регулярным) сеткам на основе шестигранников со сгущением элементов у стенок. Это позволяет уменьшить продолжительность расчета и необходимый объем оперативной памяти ЭВМ. Такие

сетки имеют выраженное сеточное направление, множество сеточных узлов упорядочено. Но построение структурной сетки зачастую является затруднительным процессом для доменов со сложной геометрической топологией. В таком случае возможно применение неструктурированной сетки, но количество элементов для одного и того же участка геометрии по сравнению с использованием структурных сеток значительно возрастет.

Конечно-объемная сетка (рис. 2) для исследуемой модели содержит 3 415 099 элементов. Точность полученного решения контролировалась по невязкам (10^{-4} для компонент скорости и турбулентных характеристик).

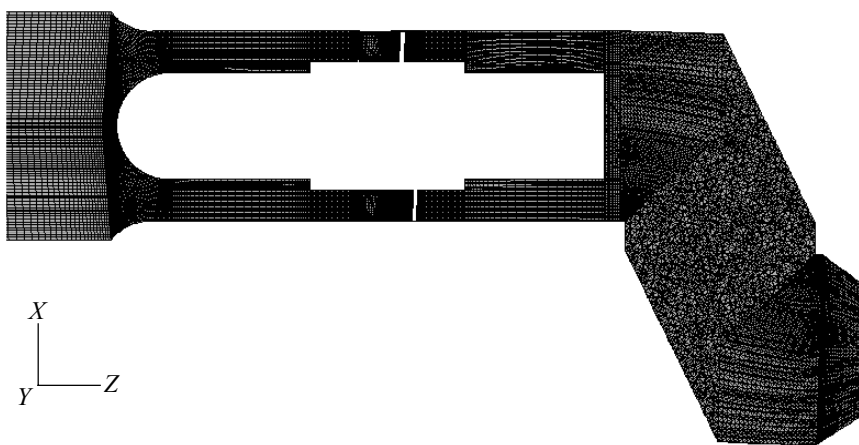


Рис. 2. Фрагмент конечно-объемной сетки в среднем сечении модели

Анализ полученных результатов расчета

Для качественного анализа течения численные результаты представляются в виде полей скоростей и статического давления в сечениях модели. На рис. 3 показана общая картина течения потока воздуха в среднем сечении проточной части модели. Перепад полного давления между двумя сечениями, указанными на рис. 4, развиваемого вентилятором при его работе на сеть составил 3085 Па.

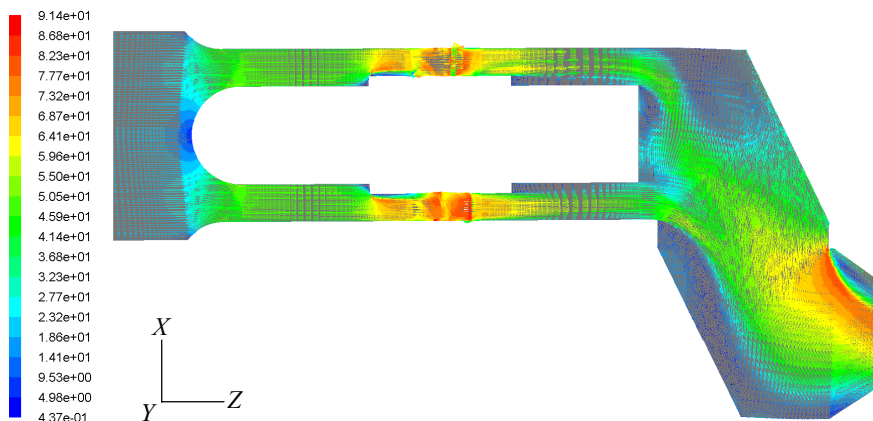


Рис. 3. Поле векторов скорости в продольном сечении модели, м/с

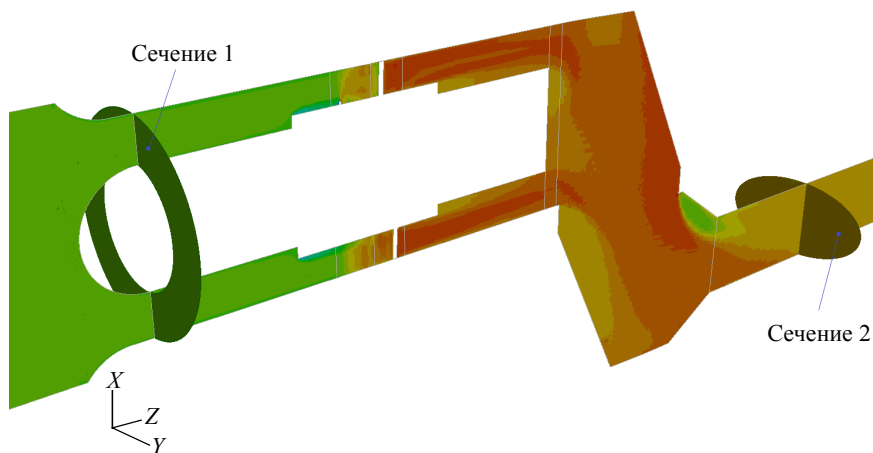


Рис. 4. Распределение избыточного полного давления в продольном сечении модели, Па

Заключение

Зачастую при увеличении сопротивления сети (разрастающаяся топология) проектировщики с целью увеличения совокупного расхода воздуха и давления потока используют несколько лопаточных машин в проекте, которые могут работать как одновременно, так и дублировать

друг друга на случай аварийной ситуации. На рис. 5 представлено поле скорости для продольного сечения модели с двумя установленными осевыми вентиляторами. В данном случае один из агрегатов является дублирующим.

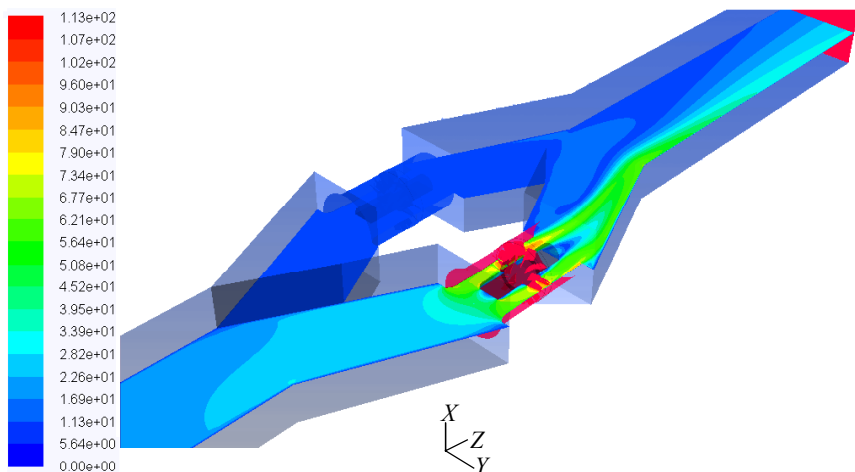


Рис. 5. Поле скорости в продольном сечении модели, м/с

На основании рассматриваемой геометрии проточной части с построенной конечно-объемной сеткой, а также с помощью программного комплекса Ansys Fluent был проведен расчет совокупной подачи воздуха, которая составила величину $Q_{v_General} = 19.15 \text{ м}^3/\text{с}$, при одновременной работе двух вентиляторов на сеть.

СПИСОК ЛИТЕРАТУРЫ

1. Волков К.Н., Емельянов В.Н. Моделирование крупных вихрей в расчетах турбулентных течений. – М.: ФИЗМАТЛИТ, 2008. – 268 с.
2. Колмогоров А.Н. Локальная структура турбулентности в несжимаемой жидкости при очень больших числах Рейнольдса // ДАН СССР. – 1941. – Т. 30. – С. 299–303.
3. Launder B.E., Spalding D.B. The Numerical computation of turbulent flows // Computer Methods in Applied Mechanics and Engineering. – 1974. – V. 3. – P. 269–289.
4. Головизнин В.М., Зайцев М.А., Карабасов С.А., Короткин И.А. Новые алгоритмы вычислительной газовой гидродинамики для многопроцессорных вычислительных комплексов. – М.: Изд-во МГУ, 2013.

5. *Гурина Е.И.* Моделирование работы шахтного вентилятора встречного вращения с помощью программного комплекса Fluent // ИФЖ. – Минск: Наука и техника, 2010. – Т. 83. – № 5. – С. 985–990.
6. *Степанов А.И.* Центробежные и осевые компрессоры, воздуходувки и вентиляторы. Теория, конструкция и применение: пер. с англ. / ред. К.З. Ушакова, А.И. Степанов. – М.: Машгиз, 1960.
7. *Идельчик И.Е.* Справочник по гидравлическим сопротивлениям. – М.: Машиностроение, 1975.
8. *Старченко А.В., Беликов Д.А., Гольдин В.Д., Нутерман Р.Б.* Пакет прикладных программ FLUENT для решения задач механики жидкости и газа, тепло и массопереноса. – Томск : Изд-во Том. унта, 2008.

Гурина Елена Ивановна, канд. физ.-мат. наук, доцент кафедры вычислительной математики и компьютерного моделирования ММФ НИ ТГУ, инженер регионального научно-образовательного математического центра НИ ТГУ; elena.gygina@mail.tsu.ru.

Исследование математической модели типа «хищник – жертва» с учетом поискового поведения хищника

А.Н. Хамидов, М.Д. Михайлов

Томский государственный университет, г. Томск, Россия

Рассматриваются два биологических вида, совместно обитающих в изолированной среде. Среда стационарна и обеспечивает в достаточной мере необходимым для жизни обоих видов ресурсом. В качестве одного из них выступает вид, называемый жертвой, а потребителем является хищник. Пространственная модель сообщества описывается системой трех дифференциальных уравнений в частных производных с соответствующими начальными и граничными условиями. Особенностью рассматриваемой модели является наличие трофической функции Холлинга $g(N) = \frac{aN}{1 + ahN}$, учитывающей насыщение рациона хищника при увеличении плотности популяции жертв, и уравнения, описывающего поисковое поведение хищника.

Ключевые слова: *модель, хищник, жертва, трофическая функция, плотность популяции, разностная схема, аппроксимация, устойчивость, сходимост, численный метод.*

Изменение плотности трофического сообщества и поисковое поведение популяции хищников описывается следующей системой дифференциальных уравнений [1]:

$$\begin{cases} \frac{\partial N}{\partial t} = rN \left(1 - \frac{N}{K}\right) - \frac{aN P}{1 + ahN} + \delta_N \frac{\partial^2 N}{\partial x^2}, \\ \frac{\partial P}{\partial t} = e \frac{aN P}{1 + ahN} - mP - \frac{\partial(Pv)}{\partial x} + \delta_P \frac{\partial^2 P}{\partial x^2}, \\ \frac{\partial v}{\partial t} = k \frac{\partial N}{\partial x} + \delta_v \frac{\partial^2 v}{\partial x^2}, \end{cases} \quad (1)$$

где $N(x, t)$, $P(x, t)$ – плотности популяций жертв и хищников, а $v(x, t)$ – скорость перемещения потребителей.

Ареал обитания представляет собой отрезок $[0, L]$ вещественной оси. Начальные условия для системы уравнений (1) имеют вид

$$\begin{cases} N(x, 0) = g_1(x), \\ P(x, 0) = g_2(x), \\ v(x, 0) = 0, \\ 0 \leq x \leq L. \end{cases} \quad (2)$$

Здесь $g_1(x)$, $g_2(x)$ – начальные функции распределения плотности популяций двух видов [2].

Граничные условия в рассматриваемой области задаются следующим образом:

$$\left. \frac{\partial N}{\partial x} \right|_{x=0, L} = \left. \frac{\partial P}{\partial x} \right|_{x=0, L} = v|_{x=0, L} = 0. \quad (3)$$

Для дальнейшего исследования поставленной задачи обезразмерим ее вводя новые переменные:

$$\begin{aligned} \tilde{x} &= \sqrt{r}x, \quad \tilde{t} = rt, \quad \tilde{N} = \frac{N}{K}, \quad \tilde{P} = \frac{P}{eK}, \quad \tilde{v} = \frac{v}{\sqrt{r}}, \\ \tilde{a} &= \frac{aeK}{r}, \quad \tilde{h} = \frac{rh}{e}, \quad \tilde{m} = \frac{m}{r}, \quad \tilde{k} = \frac{Kk}{r}, \quad \tilde{L} = \sqrt{r}L. \end{aligned}$$

После перехода к безразмерным величинам знак « \sim » над ними для простоты записи опускаем и система (1) примет вид

$$\begin{cases} \frac{\partial N}{\partial t} = N(1-N) - \frac{aN P}{1+ahN} + \delta_N \frac{\partial^2 N}{\partial x^2}, \\ \frac{\partial P}{\partial t} = \frac{aN P}{1+ahN} - mP - \frac{\partial(Pv)}{\partial x} + \delta_P \frac{\partial^2 P}{\partial x^2}, \\ \frac{\partial v}{\partial t} = k \frac{\partial N}{\partial x} + \delta_v \frac{\partial^2 v}{\partial x^2}, \end{cases} \quad (4)$$

начальные и граничные условия остаются без изменения.

Ищется решение $N, P, v \in C_1^2(G) \cap C_0^1(\bar{G})$ краевой задачи (4), (2), (3),

где

$$\begin{aligned} \bar{G} &= G \cup \Gamma, \\ G &= \{(x, t) \mid a < x < b, 0 < t \leq T\}, \\ \Gamma &= \{(x, t) \mid 0 \leq x \leq L, t = 0; x = 0, x = L, 0 < t \leq T\}, \end{aligned}$$

с помощью неявного численного метода.

Для его построения в области \bar{G} вводится прямоугольная сетка:

$$\begin{aligned}\bar{\omega}_{h\tau} &= \bar{\omega}_h \times \omega_\tau, \\ \bar{\omega}_h &= \{x_j \mid x_j = j \cdot h, j = 0, \dots, n\}, \\ \omega_\tau &= \{t_k \mid t_k = k \cdot \tau, k = 0, \dots, m\},\end{aligned}$$

и дифференциальная задача (4), (2), (3) аппроксимируется разностной:

$$\left\{ \begin{aligned} \frac{N_j^k - N_j^{k-1}}{\tau} - N_j^k (1 - N_j^{k-1}) + \frac{aN_j^k P_j^{k-1}}{1 + ah_1 N_j^{k-1}} - \delta_N \frac{N_{j+1}^k - 2N_j^k + N_{j-1}^k}{h^2} &= 0, \\ \frac{P_j^k - P_j^{k-1}}{\tau} - \frac{aN_j^k P_j^{k-1}}{1 + ah_1 N_j^k} + mP_j^k + P_j^k \frac{(v_j^{k-1} - v_{j-1}^{k-1})}{h} + v_j^{k-1} \frac{(P_j^k - P_{j-1}^k)}{h} - \\ - \delta_P \frac{P_{j+1}^k - 2P_j^k + P_{j-1}^k}{h^2} &= 0, \\ \frac{v_j^k - v_j^{k-1}}{\tau} - k_1 \frac{N_j^k - N_{j-1}^k}{h} - \delta_P \frac{v_{j+1}^k - 2v_j^k + v_{j-1}^k}{h^2} &= 0, \end{aligned} \right. \quad (5)$$

$$j = \overline{1, n-1}; \quad k = \overline{1, m};$$

$$N_j^0 = g_1(x_j), \quad P_j^0 = g_2(x_j), \quad v_j^0 = 0, \quad j = \overline{0, n}; \quad (6)$$

$$\left\{ \begin{aligned} N_0^k &= N_1^k, \quad N_n^k = N_{n-1}^k, \\ P_0^k &= P_1^k, \quad P_n^k = P_{n-1}^k, \\ v_0^k &= 0, \quad v_n^k = 0, \end{aligned} \right. \quad (7)$$

$$k = \overline{1, m}.$$

Рассмотрим вопрос о сходимости решения разностной задачи к решению дифференциальной. Для этого оценим порядок погрешности аппроксимации и покажем, что разностная схема устойчива по начальным данным. Начнем с исследования устойчивости схемы методом «замораживания коэффициентов» с однородной правой частью и однородными граничными условиями

$$\begin{cases} \frac{N_j^k - N_j^{k-1}}{\tau} - \delta_N \frac{N_{j+1}^k - 2N_j^k + N_{j-1}^k}{h^2} = 0, \\ \frac{P_j^k - P_j^{k-1}}{\tau} + v_0 \frac{P_j^k - P_{j-1}^k}{h} - \delta_P \frac{P_{j+1}^k - 2P_j^k + P_{j-1}^k}{h^2} = 0, \\ \frac{v_j^k - v_j^{k-1}}{\tau} - k_1 \frac{N_j^k - N_{j-1}^k}{h} - \delta_P \frac{v_{j+1}^k - 2v_j^k + v_{j-1}^k}{h^2} = 0, \end{cases}$$

$$j = \overline{1, n-1}; \quad k = \overline{1, m}; \quad (8)$$

$$N_j^0 = g_1(x_j), \quad P_j^0 = g_2(x_j), \quad v_j^0 = 0, \quad j = \overline{0, n}, \quad (9)$$

методом гармоник [3]. Отметим, что v_0 – это «замороженный коэффициент» v_j^{k-1} . Решение ищется в виде

$$N_j^k = \lambda_1^k e^{ij\varphi}, \quad P_j^k = \lambda_2^k e^{ij\varphi}, \quad v_j^k = \lambda_3^k e^{ij\varphi}, \quad i = \sqrt{-1}, \quad \varphi = [0, 2\pi]. \quad (10)$$

Используя обозначения

$$r_1 = \frac{\tau\delta_N}{h^2}, \quad r_2 = \frac{\tau}{h}, \quad r_3 = \frac{\tau\delta_P}{h^2}, \quad r_4 = \frac{\tau k_1}{h}, \quad r_5 = \frac{\tau\delta_v}{h^2},$$

перепишем систему (8) следующим образом:

$$\begin{cases} -r_1 N_{j-1}^k + (1 + 2r_1) N_j^k - r_1 N_{j+1}^k = N_j^{k-1}, \\ (-r_2 v_0 - r_3) P_{j-1}^k + (1 + r_2 v_0 + 2r_3) P_j^k - r_3 P_{j+1}^k = P_j^{k-1}, \\ -r_5 v_{j-1}^k + (1 + 2r_5) v_j^k - r_5 v_{j+1}^k = v_j^{k-1} + r_4 (N_j^k - N_{j-1}^k). \end{cases} \quad (11)$$

После подстановки (10) в (11) получим

$$\begin{cases} \left(1 + 4r_1 \sin^2 \frac{\varphi}{2}\right) \lambda_1^k = \lambda_1^{k-1}, \\ \left(1 + 2r_2 v_0 \sin^2 \frac{\varphi}{2} + 4r_3 \sin^2 \frac{\varphi}{2} + i r_2 v_0 \sin \varphi\right) \lambda_2^k = \lambda_2^{k-1}, \\ -r_4 \left(2 \sin^2 \frac{\varphi}{2} + i \sin \varphi\right) \lambda_1^k + \left(1 + 4r_5 \sin^2 \frac{\varphi}{2}\right) \lambda_3^k = \lambda_3^{k-1}. \end{cases} \quad (12)$$

Введем новое обозначение

$$\bar{\Lambda} = (\lambda_1, \lambda_2, \lambda_3)^T.$$

Тогда (12) переписывается в виде

$$A\bar{\Lambda}^{-k} = E\bar{\Lambda}^{-k-1}, \tag{13}$$

где

$$A = \begin{pmatrix} 1 + 4r_1 \sin^2 \frac{\varphi}{2} & 0 & 0 \\ 0 & 1 + 2r_2 v_0 \sin^2 \frac{\varphi}{2} + \\ & + 4r_3 \sin^2 \frac{\varphi}{2} + ir_2 v_0 \sin \varphi & 0 \\ -r_4 \left(2 \sin^2 \frac{\varphi}{2} + i \sin \varphi \right) & 0 & 1 + 4r_5 \sin^2 \frac{\varphi}{2} \end{pmatrix},$$

а E – единичная матрица. Равенство (13) можно представить как

$$A^{-1}A\bar{\Lambda}^{-k} = A^{-1}E\bar{\Lambda}^{-k-1}, \quad \bar{\Lambda}^{-k} = G\bar{\Lambda}^{-k-1},$$

где $G = A^{-1}$, т.е.

$$G = \begin{pmatrix} \frac{1}{1 + 4r_1 \sin^2 \frac{\varphi}{2}} & 0 & 0 \\ 0 & \boxed{1 + 2r_2 v_0 \sin^2 \frac{\varphi}{2} + \\ & + 4r_3 \sin^2 \frac{\varphi}{2} + ir_2 v_0 \sin \varphi} & 0 \\ \frac{r_4 \left(2 \sin^2 \frac{\varphi}{2} + i \sin \varphi \right)}{\left(1 + 4r_1 \sin^2 \frac{\varphi}{2} \right) \left(1 + 4r_5 \sin^2 \frac{\varphi}{2} \right)} & 0 & \frac{1}{1 + 4r_5 \sin^2 \frac{\varphi}{2}} \end{pmatrix}.$$

Запишем характеристическое уравнение матрицы G .

$$\det \begin{vmatrix} \frac{1}{1+4r_1 \sin^2 \frac{\varphi}{2}} - \mu & 0 & 0 \\ 0 & \left[\frac{1}{1+2r_2 v_0 \sin^2 \frac{\varphi}{2} + 4r_3 \sin^2 \frac{\varphi}{2} + ir_2 v_0 \sin \varphi} \right] - \mu & 0 \\ r_4 \left(2 \sin^2 \frac{\varphi}{2} + i \sin \varphi \right) & 0 & \frac{1}{1+4r_5 \sin^2 \frac{\varphi}{2}} - \mu \end{vmatrix} = 0.$$

Раскроем определитель

$$\left(\frac{1}{1+4r_1 \sin^2 \frac{\varphi}{2}} - \mu \right) \left(\frac{1}{1+2r_2 v_0 \sin^2 \frac{\varphi}{2} + 4r_3 \sin^2 \frac{\varphi}{2} + ir_2 v_0 \sin \varphi} - \mu \right) \times \\ \times \left(\frac{1}{1+4r_5 \sin^2 \frac{\varphi}{2}} - \mu \right) = 0,$$

и найдем собственные значения матрицы перехода. Проверим необходимое условие устойчивости фон Неймана [3].

$$\left(\frac{1}{1+4r_1 \sin^2 \frac{\varphi}{2}} - \mu \right) = 0, \mu_1 = \frac{1}{1+4r_1 \sin^2 \frac{\varphi}{2}}.$$

Очевидно, $4r_1 \sin^2 \frac{\varphi}{2} \geq 0$. Следовательно $|\mu_1| \leq 1$.

$$\left(\frac{1}{1+2r_2 v_0 \sin^2 \frac{\varphi}{2} + 4r_3 \sin^2 \frac{\varphi}{2} + ir_2 v_0 \sin \varphi} - \mu \right) = 0,$$

$$\mu_2 = \frac{1}{1+2r_2 v_0 \sin^2 \frac{\varphi}{2} + 4r_3 \sin^2 \frac{\varphi}{2} + ir_2 v_0 \sin \varphi}.$$

Умножим числитель и знаменатель μ_2 на величину, сопряженную к знаменателю.

$$\mu_2 = \frac{1 + 2r_2 v_0 \sin^2 \frac{\varphi}{2} + 4r_3 \sin^2 \frac{\varphi}{2} - ir_2 v_0 \sin \varphi}{\left(1 + 2r_2 v_0 \sin^2 \frac{\varphi}{2} + 4r_3 \sin^2 \frac{\varphi}{2}\right)^2 + (r_2 v_0 \sin \varphi)^2}.$$

Разделим в полученном выражении μ_2 действительную и мнимую части:

$$\mu_2 = \frac{1 + 2r_2 v_0 \sin^2 \frac{\varphi}{2} + 4r_3 \sin^2 \frac{\varphi}{2}}{\left(1 + 2r_2 v_0 \sin^2 \frac{\varphi}{2} + 4r_3 \sin^2 \frac{\varphi}{2}\right)^2 + (r_2 v_0 \sin \varphi)^2} - i \frac{r_2 v_0 \sin \varphi}{\left(1 + 2r_2 v_0 \sin^2 \frac{\varphi}{2} + 4r_3 \sin^2 \frac{\varphi}{2}\right)^2 + (r_2 v_0 \sin \varphi)^2}.$$

Путем несложных преобразований получим

$$\mu_2 = \frac{1}{\left(1 + 2r_2 v_0 \sin^2 \frac{\varphi}{2} + 4r_3 \sin^2 \frac{\varphi}{2}\right)^2 + (r_2 v_0 \sin \varphi)^2}.$$

Проверим справедливость условия $|\mu_2| \leq 1$. Очевидно, что неравенство справедливо при любых $\tau, h > 0$.

$$\left(\frac{1}{1 + 4r_5 \sin^2 \frac{\varphi}{2}} - \mu\right) = 0, \quad \mu_3 = \frac{1}{1 + 4r_5 \sin^2 \frac{\varphi}{2}}.$$

Так как

$$4r_5 \sin^2 \frac{\varphi}{2} \geq 0, \quad |\mu_3| \leq 1.$$

Отсюда следует, что схема будет абсолютно устойчивой. Очевидно, что погрешность аппроксимации имеет первый порядок по τ и по h . По теореме Лакса [3] из аппроксимации и устойчивости следует сходимость решения разностной задачи к решению дифференциальной.

Для удобства проведения численных расчетов с использованием метода прогонки и интерпретации полученных результатов разностный аналог краевой задачи (4), (2), (3) реализуется на ПЭВМ в размерном виде.

Введем обозначения:

$$a1 = -r_1, \quad b1 = 1 - \tau r \left(1 - \frac{N_j^{k-1}}{K} \right) + \tau \frac{aP_j^{k-1}}{1 + ah_1 N_j^{k-1}} + 2r_1, \quad c1 = -r_1, \quad d1 = N_j^{k-1};$$

$$a2 = -r_2 v_j^{k-1} - r_3, \quad b2 = 1 - \tau e \frac{aN_j^k}{1 + ah_1 N_j^k} + \tau m + r_2 (v_j^{k-1} - v_{j-1}^{k-1}) + r_2 v_j^{k-1} + 2r_3,$$

$$c2 = -r_3, \quad d2 = P_j^{k-1};$$

$$a3 = -r_5, \quad b3 = 1 + 2r_5, \quad c3 = -r_5, \quad d3 = v_j^{k-1} + r_4 (N_j^k - N_{j-1}^k)$$

и запишем разностную схему, аппроксимирующую систему (1), в следующем виде:

$$\begin{cases} a1N_{j-1}^k + b1N_j^k + c1N_{j+1}^k = d1, \\ a2P_{j-1}^k + b2P_j^k + c2P_{j+1}^k = d2, \\ a3v_{j-1}^k + b3v_j^k + c3v_{j+1}^k = d3, \end{cases}$$

$$j = 1, \dots, n-1; \quad k = 1, \dots, m. \quad (14)$$

Для реализации разностной задачи (14), (6), (7) с помощью метода прогонки написана программа на языке программирования высокого уровня С++. Результаты расчетов показаны на рис. 1 – 3. Следует отметить, что на этих рисунках представлены графики осредненных по пространству плотностей популяций [1]. Время t на этих графиках измеряется в сутках, а плотности двух видов – в шт./км. В качестве входных используются данные по Томской области, Верхнекетский район [4]. В качестве хищников выбрана популяция волков в количестве 121 шт. в 2017 году, а популяция жертв – зайцы указанного района в количестве 13964 шт. за тот же год.

На рис.1 представлен график изменения плотности популяции хищников на протяжении года. В течение первых трех суток наблюдается незначительный рост плотности популяции волков, а затем наступает значительное уменьшение плотности и к двадцатым суткам ее величина достигает минимального значения. Далее плотность растет и устанавливается к моменту времени $t = 100$ сут.

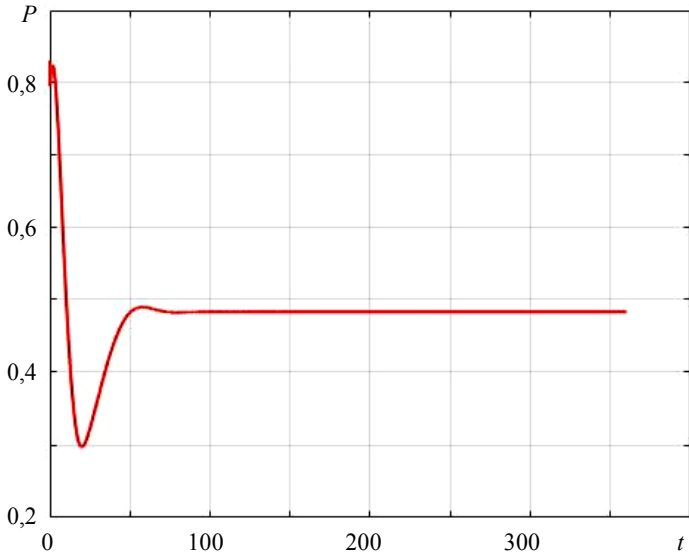


Рис. 1. График изменения плотности популяции хищников

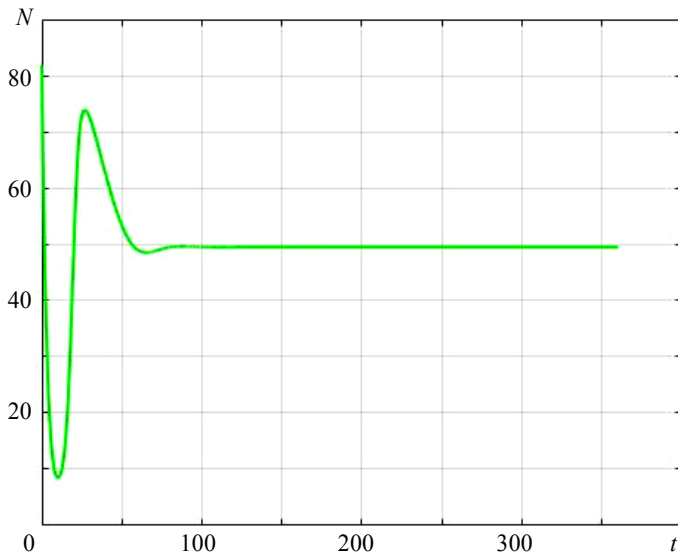


Рис. 2. График изменения плотности популяции жертв

На рис. 2 представлен график изменения плотности популяции жертв (зайцы) с течением времени t . Динамика N во многом объясняется поведением популяции хищников P и изменением скорости их перемещения v . При уменьшении плотности популяции жертв плотность популяции хищников возрастает и наоборот. Следует отметить, что, начиная с $t = 100$ сут до года, плотности популяций обоих видов сохраняют постоянные значения, что видно на рис. 1 и 2.

На рис. 3 представлена динамика поведения осредненных по пространству плотностей популяций жертв и хищников в фазовой плоскости. График демонстрирует устойчивость процессов, описываемых в рассматриваемой задаче. Фазовая траектория имеет вид спирали, которая сходится с течением времени к положению равновесия.

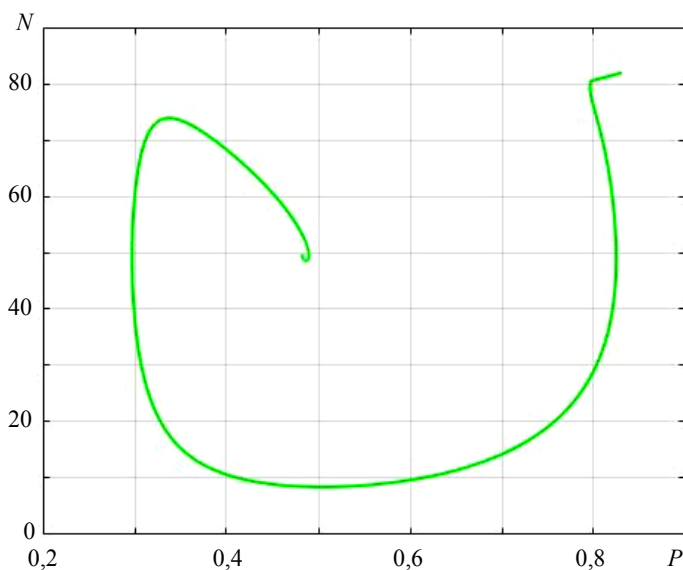


Рис. 3. Динамика поведения плотностей популяций двух видов в фазовой плоскости

Результаты численных расчетов, полученные за год, сравнивались с данными, взятыми из [5] по Верхнекетскому району за 2018 год. Относительная погрешность оказалась равной 32%.

СПИСОК ЛИТЕРАТУРЫ

1. *Тютюнов В.Ю., Сапухина Н.Ю., Сенина И.Н., Ардита Р.* Явная модель поискового поведения хищника // Журнал общей биологии. – 2002. – Т. 63. – № 2. – С. 137–148.
2. *Самарский А.А., Михайлов А.П.* Математическое моделирование. – М.: Физматлит, 2001. – 316 с.
3. *Меркулова Н.Н., Михайлов М.Д.* Методы приближенных вычислений: учеб. пособие. – 2014. – 763 с.
4. *Государственный доклад «О состоянии и охране окружающей среды Томской области в 2017 году»* [Электронный ресурс]. – URL: <https://www.ogbu.green.tsu.ru/wp-content/uploads/2021/07/Госдоклад-2017.pdf>.
5. *Государственный доклад «О состоянии и охране окружающей среды Томской области в 2018 году»* [Электронный ресурс]. – URL: <https://www.ogbu.green.tsu.ru/wp-content/uploads/2021/07/Госдоклад-2018.pdf>.

Хамидов Абдулло Нуриллаевич, магистрант НИ ТГУ; abdullo_khamidov@mail.ru;

Михайлов Михаил Дмитриевич, ст. преподаватель НИ ТГУ; michel@math.tsu.ru.

Прогноз приземной температуры воздуха на основе модели рекуррентной нейронной сети типа LSTM

И.В. Дель^{1,2}, А.В. Старченко¹

¹Томский государственный университет, г. Томск, Россия

²Институт оптики атмосферы им. В.Е. Зуева СО РАН, г. Томск, Россия

Реализована рекуррентная нейронная сеть типа LSTM для прогнозирования метеорологических значений. На основании известного распределения метеорологических значений за несколько предыдущих дней была поставлена задача прогнозирования значений приземной температуры воздуха. Общая средняя абсолютная ошибка для всего прогноза составила 2,52 °С. Сравнение различных размеров входных данных и тестовых выборок дало удовлетворительное и довольно близкое значение MAE.

Ключевые слова: прогнозирование, нейронная сеть, RNN, LSTM.

Знание точного прогноза погоды является важной информацией для различных сфер человеческой деятельности, таких как экономическая, хозяйственная и транспортная. Например, для сельского хозяйства прогноз погоды имеет огромное значение. Это дает возможность подготовиться к проливным дождям, заморозкам или засухе. Имея информацию о том, какая погода ожидается в ближайшее время, также можно предотвратить стихийные бедствия, например осведомленность о приближении наводнения, урагана или цунами позволяет разработать план спасения людей в зоне катастрофы. Такая информация носит стратегический характер, предоставляя возможность предусмотреть многие параметры, способные повлиять на образ жизни людей. Поэтому от своевременного прогноза погоды современное общество может извлечь выгоду.

Прогностические модели все время совершенствуются, повышается точность прогнозирования различных процессов. Таким образом, тема прогнозирования метеорологических величин остается актуальной.

Целью данной работы является реализация и применение модели рекуррентной нейронной сети типа LSTM (Long Short-Term Memory – долгая краткосрочная память) для прогноза приземной температуры воздуха на 24 часа. По измеренным стационарной метеостанцией историческим значениям давления, температуры и относительной влажности, скорости и направления ветра поставлена задача предсказать значения приземной температуры воздуха на 24 часа.

Архитектура рекуррентных нейронных сетей

Разработанная модель искусственной нейронной сети (ИНС) основана на одной из архитектур нейронных сетей, называемой рекуррентной нейронной сетью типа LSTM или рекуррентной нейронной сетью с долгой краткосрочной памятью.

Рекуррентные нейронные сети – это нейронные сети с обратной связью между различными слоями нейронов. Их характерная особенность – передача сигналов с выходного или скрытого слоя во входной слой. Рекуррентная нейронная сеть может состоять из любого числа слоев [1, 2].

Для обучения рекуррентных нейронных сетей применяется алгоритм обратного распространения ошибки по времени, который является вариантом алгоритма обратного распространения ошибки, используемый для нейронных сетей прямого распространения сигнала. Данный алгоритм определяет стратегию подбора весов нейронной сети с применением градиентных методов оптимизации, и также решается задача минимизации суммарных потерь на всех данных из обучающей выборки.

Обучение сети с использованием алгоритма обратного распространения ошибки проходит следующим образом. На каждом шаге обучения рекуррентной нейронной сети предъявляется некоторый входной вектор x , для которого известно ожидаемое выходное значение y . Для этого входного вектора рассчитываются веса нейронов скрытых слоев, далее – веса нейронов выходного слоя.

Выходной вектор скрытого слоя рекуррентной нейронной сети на каждом i -м шаге обучения вычисляется следующим образом:

$$t_i = \sigma_t(W_t \cdot x_i + U_t \cdot t_{i-1} + b_t),$$

где σ_t – функция активации скрытого слоя; x_i – входной вектор на i -м шаге; W_t – матрица весов для входного вектора x_i ; U_t – матрица весов для нейронов скрытого слоя; b_t – вектор смещения скрытого слоя.

В простой рекуррентной нейронной сети выходной вектор вычисляется следующим образом:

$$h_i = \sigma_h(W_h \cdot t_i + b_h),$$

где W_h – матрица весов для нейронов выходного слоя; b_h – вектор весов для нейронов выходного слоя; σ_h – функция активации для выходного слоя.

На рис. 1 представлена простая рекуррентная нейронная сеть в развернутом виде.

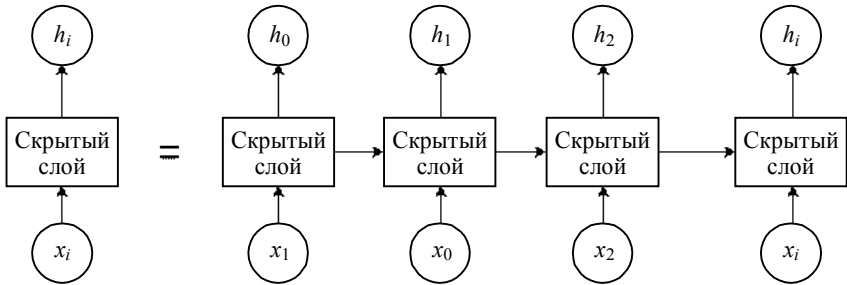


Рис. 1. Архитектура простой рекуррентной нейронной сети

Long short-term memory (LSTM) – тип рекуррентных нейронных сетей с долгой кратковременной памятью. Все рекуррентные нейронные сети представляют собой цепочки повторяющихся модулей нейронной сети. LSTM специально разработаны для устранения проблемы долгосрочной зависимости. Их специализация – запоминание информации в течение длительных периодов времени, поэтому их практически не нужно обучать [3].

Схематично архитектура LSTM-сети представлена на рис. 2, на котором каждая линия является вектором. Круг означает поточечные операции, например, суммирование векторов. Под прямоугольными

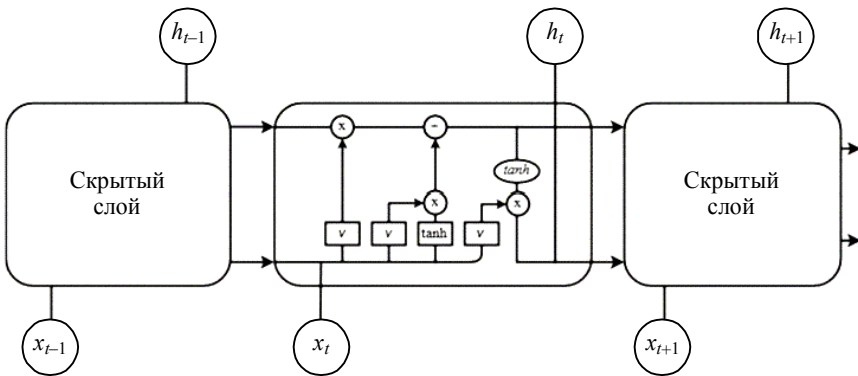


Рис. 2. Архитектура рекуррентной сети типа LSTM

ячейками понимаются слои нейронной сети (повторяющийся модуль LSTM состоит из четырех взаимодействующих слоев). Совмещение линий – это объединение векторов, а разветвления – копирование вектора с последующим хранением в разных местах.

Ключевым компонентом сети LSTM является состояние ячейки, которому сопоставляется горизонтальная линия в верхней части центрального блока. Оно участвует в некоторых линейных преобразованиях и отвечает за процесс обучения, обратное распространение ошибки и обновление весов. В LSTM можно уменьшать или увеличивать количество информации в состоянии ячейки при помощи структур, называемых гейтами. Гейт – это «ворота», либо пропускающие, либо не пропускающие информацию [3].

Гейты состоят из сигмоидального слоя (т.е. предполагают использование сигмоидальной функции активации) и операции поточечного умножения. Сигмоидальный слой возвращает число в диапазоне от «0» до «1», определяющее, какую долю каждого блока информации следует пропустить дальше по сети. Число «0» означает «не пропускать ничего», а число «1» – «пропустить все». LSTM имеет три таких гейта для контроля состояния ячейки [3].

На первом этапе сеть решает, какую информацию можно удалить из состояния ячейки. Это решение принимается сигмоидальным слоем, называемым «слоем гейта утраты». Он выдает число от 0 до 1 для каждого номера в состоянии ячейки. 1 означает «полностью сохранить», а 0 – «полностью удалить» [3].

Второй шаг – решить, какая новая информация будет храниться в состоянии ячейки. Этот этап состоит из двух частей. Сначала сигмоидальный слой, называемый «слоем гейта входа», решает, какие значения требуется обновить. Затем слой \tanh создает новый вектор состояния ячеек [3].

На третьем шаге LSTM-сеть изменяет старое состояние ячейки на новое состояние.

Наконец, нужно решить, что требуется получить на выходе. Результатом будет являться отфильтрованное состояние ячейки. Сначала используется сигмоидальный слой, который решает, какие части состояния ячейки выводить. Затем состояние ячейки пропускается через \tanh (чтобы разместить все значения в интервале $[-1, 1]$) и умножается на выходной сигнал сигмоидального гейта [3].

Построение модели рекуррентной нейронной сети типа LSTM

Задача состоит в предсказании приземной температуры воздуха на 24 часа, используя почасовую температуру за 48 часов.

Исходные метеорологические данные для исследования были получены с сайта «Расписание погоды» (<http://rp5.ru>) за период времени с 01.01.2014 по 1.11.2021. Размер набора данных – 68688 ежечасных наблюдений.

После сбора данных проводится процедура их предварительной обработки: восстановление пропусков. Затем все параметры нормализуются. Далее выбранные данные разделяются на две выборки: обучающую выборку, соответствующую 80% данных от основной выборки, и тестовую выборку, соответствующую 20%, чтобы можно было проверить способность сети к прогнозированию после фазы обучения.

Формой ввода является 48 временных шагов с 7 параметрами. Скрытый слой рекуррентной нейронной сети LSTM содержит 10 нейронов. Выходной слой – 24 нейрона [4].

В этой нейронной сети осуществляется оптимизация целевой функции с подбором весовых коэффициентов методом адаптивной инерции Adam (Adaptive Moment Estimation) (вариант стохастического градиентного спуска) [5].

Метрикой для рассматриваемой задачи в данной работе будет средняя абсолютная ошибка (MAE), рассчитываемая для каждого прогнозного временного шага.

За функцию потерь также принята MAE:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |d_i - y_i|,$$

где y_i – предсказанное выходное значение; d_i – реальное выходное значение. Функция потерь используется для расчета ошибки между реальными и полученными значениями. Глобальная цель – минимизировать эту ошибку.

Сеть обучалась за фиксированное количество эпох (epochs = 30). Количество эпох показывает, сколько раз модель подвергается воздействию обучения. Число эпох выбрано экспериментальным путем.

Близость графиков на рис. 3 свидетельствует об удовлетворительной настройке сети.

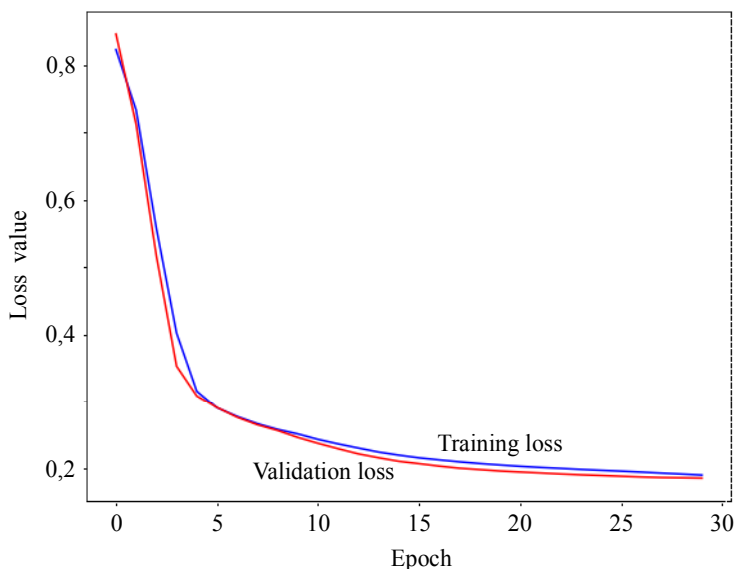


Рис. 3. Потери на этапах обучения и проверки

Результаты прогнозирования

Общее значение MAE по всему прогнозу составило 2,52 °С. Результат удовлетворительный.

В таблице приведены результаты значений MAE по выбранным случайным образом дням из тестовой выборки. При сравнении этих значений друг с другом видно, что в некоторые дни ИНС хуже, а в другие лучше справляется с прогнозированием.

Некоторые результаты значений MAE

Дата	MAE, °С
16.04.2020	1.06
15.02.2021	4.87
22.06.2020	2.45
10.09.2020	1.59
16.04.2021	4.37
19.04.2021	8.74
18.07.2020	1.35
06.08.2020	1.38

Для примера на рис. 4 – 7 представлены графики сравнения прогнозируемой почасовой температуры с наблюдениями в разные даты из тестовой выборки.

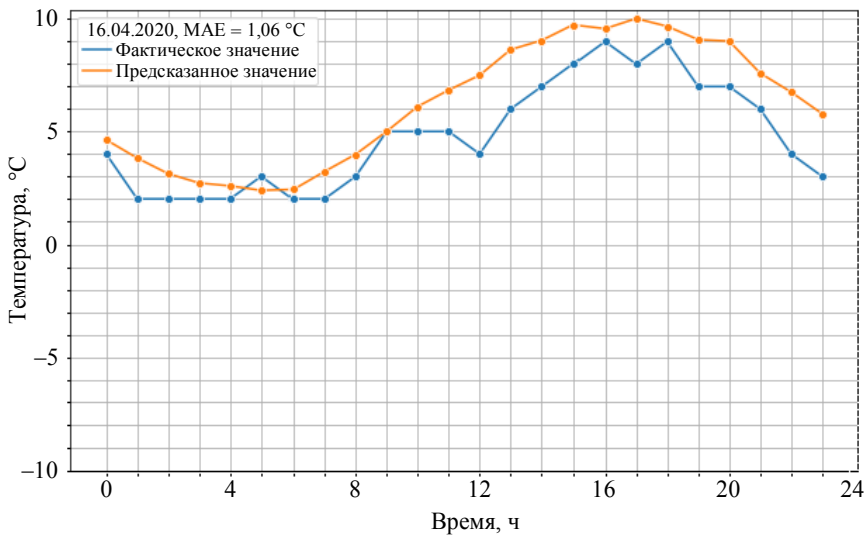


Рис. 4. Сравнение прогнозируемой почасовой температуры с наблюдениями за 16.04.2020, MAE = 1,06 °C

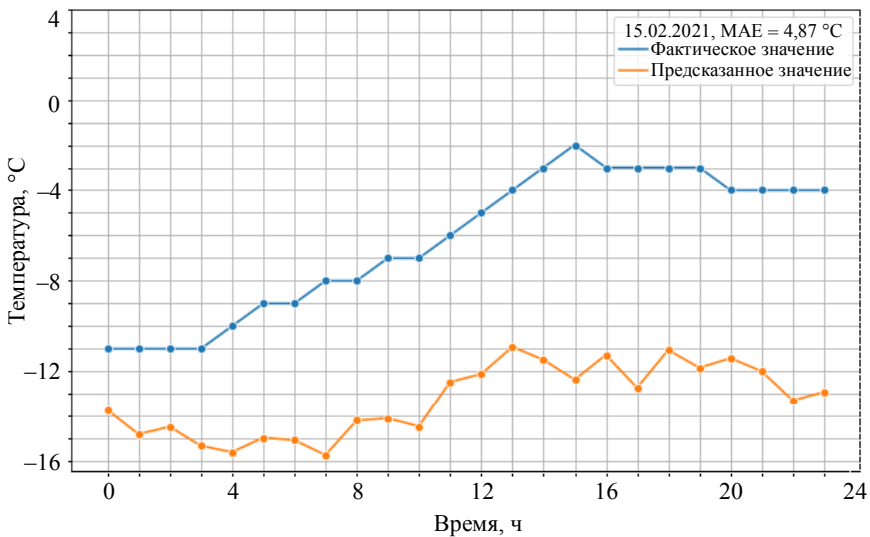


Рис. 5. Сравнение прогнозируемой почасовой температуры с наблюдениями за 15.02.2021, MAE = 4,87 °C

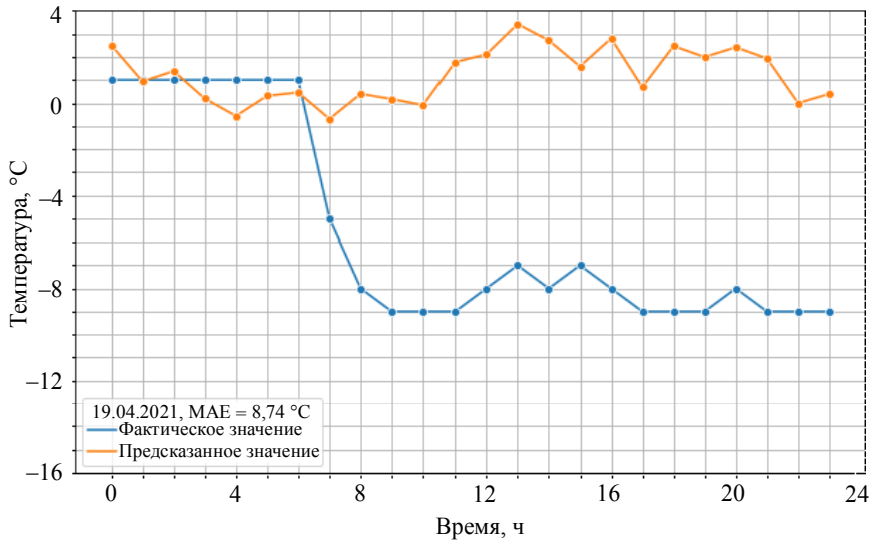


Рис. 6. Сравнение прогнозируемой почасовой температуры с наблюдениями за 19.04.2021, MAE = 8,74 °C

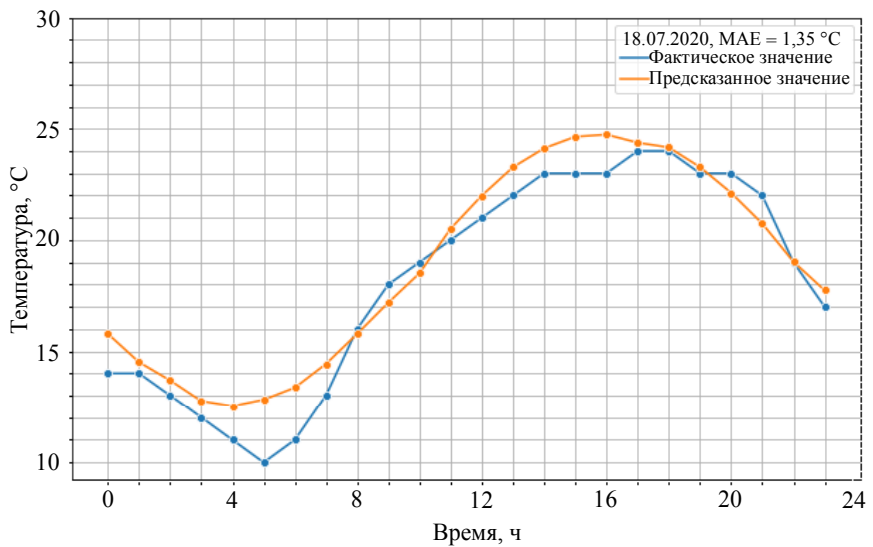


Рис. 7. Сравнение прогнозируемой почасовой температуры с наблюдениями за 18.07.2020, MAE = 1,35 °C

Также был проведен сравнительный анализ средней абсолютной ошибки при различных размерах входных данных и тестовой выборки. Сравнивая значения MAE при различном размере тестовой выборки (10, 15, 20 и 30%) и размере входных данных (24, 48, 72, 96 и 120 наблюдений), можно сделать вывод, что во всех случаях результат удовлетворительный и довольно схожий, что видно из рис. 4.

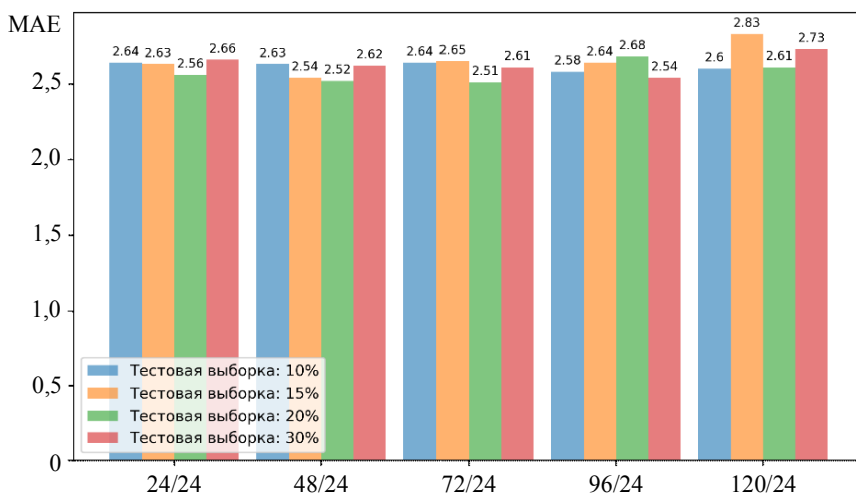


Рис. 8. Сравнение значений MAE для разных размеров входных данных и разных методов оптимизации

Заключение

В данной работе была разработана модель рекуррентной нейронной сети типа LSTM с дальнейшим ее применением для прогноза приземной температуры воздуха в г. Томске. По измеренным стационарной метеостанцией значениям давления, температуры и относительной влажности, скорости и направления ветра предсказывались значения приземной температуры воздуха на 24 часа с использованием библиотеки Keras языка программирования Python.

Сравнительный анализ реальных данных с результатами, полученными при помощи LSTM-сети, показал, что метод нейросетевого прогнозирования может конкурировать со стандартными методами прогнозирования метеорологических параметров.

СПИСОК ЛИТЕРАТУРЫ

1. *Hochreiter S., Schmidhuber J.* Long Short-Term Memory // *Neural Computation.* – 1997. – V. 9. – No. 8. – P. 1735–1780.
2. *Хайкин С.* Нейронные сети. Полный курс. – М: Вильямс, 2008. – 1104 с.
3. *Гафаров Ф.М., Галимянов А.Ф.* Искусственные нейронные сети и их приложения. – Казань: Изд-во Казан. ун-та, 2018. – 121 с.
4. *Jason Brownlee.* Deep Learning for Time Series Forecasting. Predict the Future with MLPs, CNNs and LSTMs in Python. – 2018. – 557 p.
5. *Diederik P. Kingma, Jimmy Ba. Adam.* A Method for Stochastic Optimization // *ICLR* 2015. – 2015.

Дель Ирина Васильевна, магистрант ММФ НИ ТГУ, мл. науч. сотр. ИОА СО РАН; irina.del@math.tsu.ru.

Старченко Александр Васильевич, докт. физ.-мат. наук, профессор, зав. кафедрой вычислительной математики и компьютерного моделирования НИ ТГУ; starch@math.tsu.ru;

О построении преобуславливателей для четно-нечетного расщепления сплайн-вейвлетов

Б.М. Шумилов

*Томский государственный архитектурно-строительный университет,
г. Томск, Россия*

Исследован метод разложения сплайнов степени m гладкости C^{m-1} на серию вейвлетов с нулевыми моментами и смещенными носителями. Оригинальность состоит в применении к полученной системе линейных алгебраических уравнений, связывающих коэффициенты сплайн-разложения на начальном масштабе со сплайн-коэффициентами и вейвлет-коэффициентами на вложенном масштабе, некоторого преобуславливателя, сводящего систему к последовательности явных усредняющих соотношений. Представлен пример применения метода для случая сплайнов первой степени и вейвлетов с двумя первыми нулевыми моментами. После исключения четных строк системы результирующая матрица приобретает трехдиагональный вид, что позволяет применить к ее решению эффективный метод прогонки.

Ключевые слова: *V-сплайны, вейвлеты, неявные соотношения разложения, метод прогонки.*

Вейвлетом называется короткая или быстро затухающая волновая функция (всплеск), для которой множество сжатий и смещений порождает пространство измеримых функций на всей числовой оси. Основой для построения вейвлетов является наличие набора аппроксимирующих пространств $\dots V_{L-1} \subset V_L \subset V_{L+1} \dots$ таких, что каждая базисная функция в V_L может быть выражена в виде линейной комбинации базисных функций в V_{L+1} . В частности, таким свойством обладают сплайны – гладкие функции, склеенные из кусков многочленов степени m , на вложенной последовательности сеток. Суть вейвлет-преобразования можно сформулировать следующим образом: оно позволяет разложить заданную функцию V_{L+1} на грубое приближенное представление V_L и локальные уточняющие подробности $W_L = V_{L+1} - V_L$. Главное при этом – найти подходящий базис пространства W_L и построить для него быстрые взаимно-однозначные формулы прямого и обратного вейвлет-преобразований.

Пусть пространство V_L является пространством сплайнов степени m гладкости C^{m-1} на равномерной бесконечно продолженной в обе сто-

роны сетке узлов Δ_L : $x_{i+1} = x_i + 1/2^L$, для всех i . Хорошо известно, что базис в этом пространстве порождают функции $\varphi_m(v - i) \forall i$, где $v = 2^L x$, порожденные сжатиями и сдвигами функции вида

$$\varphi_m(t) = \frac{1}{m!} \sum_{j=0}^{m+1} (-1)^j \binom{m+1}{j} (t-j)_+^m,$$

где $t_+^m = (\max\{t, 0\})^m$, и что они удовлетворяют калибровочному соотношению

$$\varphi_m(t) = 2^{-m} \sum_{k=0}^{m+1} \binom{m+1}{k} \varphi_m(2t - k). \quad (1)$$

В результате любой сплайн на сетке Δ_L может быть представлен в виде

$$S^L(x) = \sum_{-\infty}^{\infty} C_i^L \varphi_m(2^L x - i), \quad (2)$$

где коэффициенты $C_i^L \forall i$ являются решением, например, задачи кардинальной интерполяции:

$$S^L(x_i) = f(x_i), \quad -\infty < i < \infty.$$

Если сетка Δ_{L-1} получена из сетки Δ_L посредством удаления каждого второго узла, то соответствующее пространство V_{L-1} с базисными функциями $\varphi_m(v/2 - i) \forall i$, у которых носители в 2 раза больше по ширине, вложено в V_L . Разность пространств V_L и V_{L-1} составляет пространство вейвлетов $W_{L-1} = V_L - V_{L-1}$ [1].

В качестве базисных функций в пространстве W_{L-1} можно использовать ортогональные вейвлеты – элементы пространства V_L , ортогональные как всем элементам пространства V_{L-1} , так и между собой, либо полуортогональные вейвлеты – элементы пространства V_L , ортогональные всем элементам пространства V_{L-1} , но не ортогональные между собой. Разница между ортогональными и полуортогональными вейвлетами состоит в наличии либо в отсутствии явных конечных формул разложения и, наоборот, в отсутствии либо в наличии явных аналитических формул для вычисления самих вейвлетов. Общим является то, что вейвлет-разложение обеспечивает наилучшее среднеквадратическое приближение сплайнов на густой сетке посредством сплайнов на прореженной сетке и что носители вейвлетов жестко увязаны с носителями базисных сплайнов. В отличие от этого, далее мы будем использовать вейвлеты, ортогональные всем многочленам степени n [2],

$$w_{m,n}(t) = 2^{-m} \sum_{k=0}^{n+1} (-1)^k \binom{n+1}{k} \Phi_m(2t-k). \quad (3)$$

Они имеют следующие носители:

$$\text{supp } w_{m,n} = [0, n+1],$$

и соответственно $n+1$ нулевых моментов

$$\int_0^{n+1} x^k w_{m,n}(x) dx = 0,$$

для $k = 0, 1, \dots, n$. Вдобавок, их можно перемещать, в определенных пределах, относительно носителей базисных сплайнов.

Запишем базисные сплайн-функции в виде однострочной матрицы бесконечной длины

$$\varphi^L(x) = [\dots, \Phi_m(2^L x - i), \Phi_m(2^L x - i - 1), \dots].$$

Введя обозначение $\mathbf{c}^L = [\dots, C_i^L, C_{i+1}^L, \dots]^T$ для вектора, состоящего из коэффициентов сплайна, запишем формулу (2) в векторном виде $s^L(x) = \varphi^L(x) \mathbf{c}^L$.

Точно так же мы можем записать базисные вейвлет-функции на уровне разложения L в виде бесконечной матрицы-строки

$$\psi^{L-1}(x) = [\dots, w_{m,n}(2^L x - i), w_{m,n}(2^L x - i - 1), \dots].$$

Соответствующие вейвлет-коэффициенты аппроксимации обозначим через D_i^{L-1} , $-\infty < i < \infty$, и введем вектор-столбец

$$\mathbf{d}^{L-1} = [\dots, D_i^{L-1}, D_{i+1}^{L-1}, \dots]^T.$$

Поскольку пространства V_{L-1} и W_{L-1} по определению являются подпространствами V_L , функции $\varphi^{L-1}(x)$ и $\psi^{L-1}(x)$ можно представить в виде линейных комбинаций функций $\varphi^L(x)$:

$$\varphi^{L-1}(x) = \varphi^L(x) P^L \quad \text{и} \quad \psi^{L-1}(x) = \varphi^L(x) Q^L, \quad (4)$$

где столбцы матрицы P^L составлены из коэффициентов соотношений (1), поскольку каждая широкая базисная функция может быть построена из $(m+2)$ -х узких базисных функций, а элементы столбцов матрицы Q^L составлены из коэффициентов соотношений (3).

Следовательно имеет место цепочка равенств:

$$\begin{aligned}\varphi^L(x)\mathbf{c}^L &= \varphi^{L-1}(x)\mathbf{c}^{L-1} + \psi^{L-1}(x)\mathbf{d}^{L-1} = \\ &= \varphi^L(x)P^L\mathbf{c}^{L-1} + \varphi^L(x)Q^L\mathbf{d}^{L-1}.\end{aligned}\quad (5)$$

Пусть известны коэффициенты \mathbf{c}^{L-1} и \mathbf{d}^{L-1} . Тогда коэффициенты \mathbf{c}^L могут быть легко получены из \mathbf{c}^{L-1} и \mathbf{d}^{L-1} следующим образом:

$$\mathbf{c}^L = P^L\mathbf{c}^{L-1} + Q^L\mathbf{d}^{L-1}.\quad (6)$$

Формула (6) представляет собой не что иное, как алгоритм восстановления [1, с. 248], для реализации которого в силу ленточности матриц P^L и Q^L с успехом используется схема скользящего среднего. К сожалению, для обратного процесса вычисления из коэффициентов \mathbf{c}^L более грубой версии $\mathbf{c}^{L-1} = A^L\mathbf{c}^L$ и уточняющих коэффициентов $\mathbf{d}^{L-1} = B^L\mathbf{d}^L$, следуя алгоритму разложения [1, с. 247], для всех случаев, кроме случая вейвлетов Хаара, получаем, что строки матриц A^L и B^L – это бесконечные числовые последовательности, и их усечение приводит к неизбежным погрешностям.

Используя обозначения для блочных матриц, перепишем равенство (6) в виде

$$\mathbf{c}^L = \left[P^L \mid Q^L \right] \begin{bmatrix} \mathbf{c}^{L-1} \\ \mathbf{d}^{L-1} \end{bmatrix}\quad (7)$$

и определим некоторые бесконечные матрицы R^L и G^L условиями:

а) $G^L = \left[P^L \mid Q^L \right] R^L$;

б) матрица R^L – ленточная, с минимально возможным количеством ненулевых диагоналей;

в) матрица G^L – ленточная, с ненулевыми элементами на главной диагонали и на нечетных позициях нечетных столбцов.

Тогда, предполагая невырожденность матрицы G^L , домножим левую и правую стороны равенства (7) на матрицу $R^L G^{L-1}$. В результате получим

$$\begin{aligned}R^L G^{L-1} \mathbf{c}^L &= R^L \left(\left[P^L \mid Q^L \right] R^L \right)^{-1} \left[P^L \mid Q^L \right] \begin{bmatrix} \mathbf{c}^{L-1} \\ \mathbf{d}^{L-1} \end{bmatrix} = \\ &= R^L R^{L-1} \left[P^L \mid Q^L \right]^{-1} \left[P^L \mid Q^L \right] \begin{bmatrix} \mathbf{c}^{L-1} \\ \mathbf{d}^{L-1} \end{bmatrix} = \begin{bmatrix} \mathbf{c}^{L-1} \\ \mathbf{d}^{L-1} \end{bmatrix}.\end{aligned}$$

Таким образом, вместо того, чтобы пытаться непосредственно решать систему вида (7), мы можем решить систему

$$G^L \xi^L = \mathbf{c}^L \quad (8)$$

относительно некоторых значений ξ^L и после этого просто вычислить значения \mathbf{c}^{L-1} и \mathbf{d}^{L-1} с использованием линейного преобразования

$$\begin{bmatrix} \mathbf{c}^{L-1} \\ \mathbf{d}^{L-1} \end{bmatrix} = R^L \xi^L.$$

С целью иллюстрации применения данной технологии рассмотрим далее конкретный пример.

Сплайн-вейвлеты 1-й степени с 2 нулевыми моментами. Для случая $m = 1$ базисные сплайн-функции принимают особенно простой вид (рис. 1)

$$\varphi_1(t) = \begin{cases} t, & \text{если } 0 \leq t \leq 1; \\ 2-t, & \text{если } 1 \leq t \leq 2; \\ 0, & \text{иначе.} \end{cases} \quad (9)$$

График вейвлета 1-й степени, ортогонального всем многочленам 1-й степени, изображен на рис. 2.

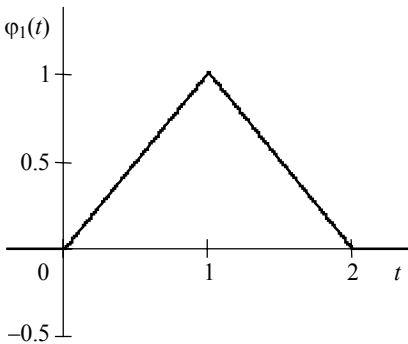


Рис. 1. График масштабировющей функции $\varphi_1(t)$

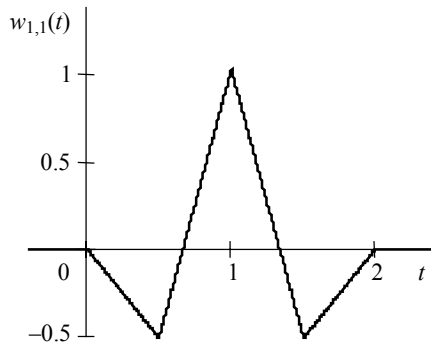


Рис. 2. График функции $w_{1,1}(t)$

Хорошо известно, что центры носителей полуортогональных вейвлетов 1-й степени сдвинуты на полшага относительно центров носителей базисных сплайнов, что приводит при обращении к полностью заполненным матрицам. То же самое справедливо и для так называемых неортогональных вейвлетов [2]. Пусть теперь центры носителей вейвлетов и базисных сплайнов совпадают. В результате бесконечная матрица $[P^L | Q^L]$ приобретает следующий вид:

$$[P^L | Q^L] = \frac{1}{2} \left[\begin{array}{cccc|cccc} \vdots & & & & \vdots & & & \\ \ddots & 0 & & & \ddots & 0 & & \\ \ddots & 1 & \vdots & & \ddots & 1 & \vdots & \\ \ddots & 2 & 0 & & \ddots & -2 & 0 & \\ & 1 & 1 & \ddots & & 1 & 1 & \ddots \\ & 0 & 2 & \ddots & & 0 & -2 & \ddots \\ & \vdots & 1 & \ddots & & \vdots & 1 & \ddots \\ & & 0 & & & & 0 & \ddots \\ & & \vdots & & & & \vdots & \end{array} \right].$$

Для матрицы $[P^L | Q^L]$ четырехдиагональная матрица

$$[P^L | Q^L]' = \frac{1}{2} \left[\begin{array}{cccccc|cccc} \ddots & \ddots & \vdots & \vdots & & & & & & \\ \ddots & -2 & 0 & 0 & \vdots & & & & & \\ \ddots & 1 & 1 & 1 & 0 & \vdots & & & & \\ \cdots & 0 & 2 & -2 & 0 & 0 & \ddots & & & \\ \cdots & 0 & 1 & 1 & 1 & 1 & 0 & \cdots & & \\ & \ddots & 0 & 0 & 2 & -2 & 0 & \cdots & & \\ & & \vdots & 0 & 1 & 1 & 1 & \ddots & & \\ & & & \vdots & 0 & 0 & 2 & \ddots & & \\ & & & & \vdots & \vdots & \ddots & \ddots & & \end{array} \right]$$

получается перестановкой столбцов матрицы $[P^L | Q^L]$ так, чтобы столбцы матриц P^L и Q^L перемежались. На практике такая перестановка сопровождается изменением порядка неизвестных в системе (7) и часто делается с целью придания системе ленточного вида для облегчения численного решения системы [3]. Однако, как видно из приведенного выше примера, полученная система уравнений не имеет диагонального

преобладания, что может затруднить вейвлет-анализ данных большого размера.

Матрицу, соответствующую указанной перестановке столбцов, обозначим через T . Тогда верно представление [4]

$$\left[P^L \mid Q^L \right]' = \left[P^L \mid Q^L \right] T. \tag{10}$$

Из представления (10) находим

$$\left[P^L \mid Q^L \right]^{-1} G^L = T \left[P^L \mid Q^L \right]'^{-1} G^L. \tag{11}$$

Таким образом, проблема свелась к отысканию решения системы матричных равенств

$$\left[P^L \mid Q^L \right]'_j R'^L_j = G^L_j, \quad -\infty < j < \infty. \tag{12}$$

Здесь нижние индексы в обозначениях матриц указывают на то, для вычисления элементов каких столбцов матрицы R'^L предназначена соответствующая система (12). При этом, помещая в верхнюю и нижнюю части каждого столбца матрицы R'^L столько нулей, сколько представляется возможным, мы обеспечиваем компактность вычислительной схемы; обнуляя элементы, отстоящие на нечетное количество шагов от главной диагонали матрицы G^L , мы обеспечиваем возможность использования для решения эффективного метода прогонки; а дополнительно обнуляя элементы, стоящие вне главной диагонали матрицы G^L , мы обеспечиваем возможность последующего расщепления системы на четные и нечетные строки.

Конкретно, согласно предполагаемой ступенчатой структуре матриц R'^L и G^L , система (12) распадается на блоки с матрицами следующего вида:

$$\left[P^L \mid Q^L \right]'_{j,j+1} = \frac{1}{2} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 2 & -2 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 2 & -2 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

при условии, что уравнения, соответствующие нулевым строкам матрицы G^L_j удалены из системы. Данная система совместная и недоопреде-

3) Значения ξ^L в четных узлах приравниваются соответствующим коэффициентам сплайна:

$$\xi_i^L = C_i^L, \quad i = 0, 2, \dots, 2^L.$$

4) Значения сплайн-коэффициентов на прореженной сетке вычисляются по формулам

$$\begin{aligned} C_0^{L-1} &= C_1^L + \frac{1}{2}C_0^L, \\ C_i^{L-1} &= C_{2i-1}^L + \frac{1}{2}C_{2i}^L + C_{2i+1}^L, \\ i &= 1, 2, \dots, 2^{L-1}, \end{aligned}$$

тогда как вейвлет-коэффициенты равны

$$\begin{aligned} D_0^{L-1} &= C_1^L - \frac{1}{2}C_0^L, \\ D_i^{L-1} &= C_{2i-1}^L - \frac{1}{2}C_{2i}^L + C_{2i+1}^L, \\ i &= 1, 2, \dots, 2^{L-1}. \end{aligned}$$

5) Если $L \geq 2$, то значение L уменьшается на 1, и осуществляется переход к шагу 2.

6) Иначе, на всех уровнях разложения L выполняется отбраковка незначимых вейвлет-коэффициентов согласно некоторому критерию [7], и последовательно восстанавливаются коэффициенты сплайна согласно алгоритму скользящего среднего (6).

7) После восстановления (безусловно, приближенного) коэффициентов сплайна для самой густой сетки к ним добавляется вычтенное ранее на 1-м этапе значение измерения на правом конце:

$$f_i = C_i^L + f_{2^L}, \quad i = 0, 1, \dots, 2^L.$$

Результаты численных экспериментов по предложенному алгоритму были представлены ранее в работах [5, 8]. А в работе [9] был исследован аналогичный метод разложения кубических сплайнов гладкости C^2 на серию вейвлетов с нулевыми моментами и смещенными носителями. Для случая шести нулевых моментов обнаружено отсутствие строгого диагонального преобладания, однако для семидиагональной реализации выполнена модификация крайних уравнений, обеспечивающая строгое диагональное преобладание по столбцам системы и, как следствие, устойчивость вычислений.

СПИСОК ЛИТЕРАТУРЫ

1. Чуи Ч. Введение в вейвлеты: пер. с англ. – М.: Мир, 2001. – 412 с.
2. Koro K., Abe K. Non-orthogonal spline wavelets for boundary element analysis // Engineering Analysis with Boundary Elements. – 2001. – V. 25. – P. 149–164.
3. Столниц Э., ДеРоуз Т., Салезин Д. Вейвлеты в компьютерной графике: пер. с англ. – Ижевск: НИЦ «Регулярная и хаотическая динамика», 2002. – 272 с.
4. Писсанецки С. Технология разреженных матриц: пер. с англ. – М.: Мир, 1988. – 410 с.
5. Шумилов Б.М. Алгоритмы с расщеплением вейвлет-преобразования сплайнов первой степени на неравномерных сетках // Журнал вычислительной математики и математической физики. – 2016. – Т. 56. – № 7. – С. 39–50.
6. Самарский А.А., Николаев Е.С. Методы решения сеточных уравнений. – М.: Наука, 1978. – 592 с.
7. Воробьев В.И., Грибунин В.Г. Теория и практика вейвлет-преобразования. – СПб.: ВУС, 1999. – 202 с.
8. Шумилов Б.М., Матанов Ш.М. Алгоритм с расщеплением вейвлет-преобразования сплайнов первой степени // Вестн. Том. гос. ун-та. Управление, вычисл. техника и информатика. – 2011. – № 3 (16). – С. 51–57.
9. Шумилов Б.М. О расщеплении для кубических сплайн-вейвлетов с четырьмя нулевыми моментами на отрезке // Вычислительные технологии. – 2021. – Т. 26. – № 2. – С. 72–87.

Шумилов Борис Михайлович, докт. физ.-мат. наук, профессор кафедры прикладной математики ТГАСУ; b_shumilov@math.tsu.ru.

О кратномасштабном вейвлет-анализе сеточных функций

В.Н. Берцун

Томский государственный университет, г. Томск, Россия

Термин «вейвлет» в переводе с английского означает «маленькая волна» или «короткий всплеск». Его ввели Гроссман и Морле в середине 80-х годов в связи с анализом свойств сейсмических и акустических сигналов [3, 11]. Вейвлет-анализ является одним из востребованных разделов математики и ее приложений как составная часть технологии мягких вычислений, используемая для предварительной обработки данных в системах искусственного интеллекта. Кратномасштабный вейвлет-анализ, который основан на использовании последовательности вложенных подпространств, позволяет анализировать сигнал одновременно для разных масштабов [5]. Основной идеей этой теории является разделение исходного сигнала на два потока A и B, один из которых (A) отображает средние значения, а другой (B) напротив – быстроменяющуюся часть потока. Рассматриваются свойства ортонормированных масштабирующих функций и вейвлетов Хаара. Приводятся примеры кратномасштабного анализа сигналов.

Ключевые слова: *цифровой сигнал, вейвлеты Хаара, кардинальные сплайны.*

Любая упорядоченная последовательность данных, заданная аналитически или в виде сеточной функции, называется сигналом. Совершенствование и применение методов цифровой обработки сигналов и изображений имеет важное значение для решения различных научно-технических задач. На практике сигнал всегда содержит не только полезную информацию, но и следы некоторых внешних воздействий, что вносит дополнительную сложность в процесс обработки данных. Термин «вейвлет» в переводе с английского означает «маленькая волна» или «короткий всплеск» [1–3]. Вейвлет-анализ применяется для нестационарных сигналов, связанных с широким спектром прикладных задач: квантовой механики, обработки изображений и сжатия данных, компьютерной безопасности и анализа кровяного давления, исследования климата и анализа ДНК, предсказания и предотвращения возможных катастроф, очистки данных в системах искусственного интеллекта и др. Такой анализ существенно использует достижения в области теории сплайнов и преобразования Фурье, позволяет выделять из сигнала

компоненты разного масштаба. Основной идеей кратномасштабного вейвлет-анализа (КМА), предложенного в 1988–1989 гг. Малла и И. Мейром, является разделение исходного потока информации на два потока, один из которых отображает средние значения, а другой – быстроменяющуюся часть потока.

В процессе вейвлет-преобразования (ВП) исходный дискретный сигнал раскладывается на высокочастотные детали и сглаженную уменьшенную версию оригинала. Это достигается применением двух фильтров, причем каждая из полученных составляющих вдвое меньше исходного сигнала.

Освоение, совершенствование и применение математического моделирования структуры сложных нестационарных сигналов с помощью вейвлет-преобразования являются составной частью методологии мягких вычислений и вычислительного интеллекта [4]. ВП одномерного сигнала – это его представление в виде обобщенного ряда или интеграла Фурье по системе базисных функций [5]

$$\Psi_{ab}(x) = \frac{1}{\sqrt{a}} \Psi\left(\frac{x-b}{a}\right), \quad a, b \in R, \quad a > 0,$$

где $\Psi(x)$ – материнский вейвлет; a – параметр масштаба; b – параметр сдвига по оси Ox . Малые значения параметра a соответствуют высоким частотам, большие – малым.

Непрерывное ВП для $f(x) \in L_2(R)$ записывается в виде

$$[W_\Psi f](a, b) = (f, \Psi_{a,b}) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(x) \overline{\Psi\left(\frac{x-b}{a}\right)} dx,$$

а обратное определяется по формуле

$$f(x) = \frac{1}{C_\Psi} \iint [W_\Psi f](a, b) \Psi_{a,b}(x) \frac{dadb}{a^2},$$

где $C_\Psi = \int_{-\infty}^{\infty} |\widehat{\Psi}(\omega)|^2 |\omega|^{-1} d\omega < \infty$, $\widehat{\Psi}(\omega) = \int_{-\infty}^{\infty} \Psi(x) e^{-i\omega x} dx$, $\widehat{\Psi}(0) = 0$.

В качестве базисных функций, образующих ортогональный базис, используется широкий набор вейвлетов. Вейвлет-функция должна обладать следующими признаками: локализация, нулевое среднее значение момента, ограниченность и автомодельность базиса [6]. Такими свойствами обладают, например, вейвлеты, полученные на основе функции Гаусса n -го порядка, представленные на рис. 1.

1. WAVE-вейвлет первого порядка

$$\psi(x) = -x \exp\left(-\frac{x^2}{2}\right),$$

$$\psi_{ab}(x) = \frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a}\right) = -\frac{1}{\sqrt{a}} \left(\frac{x-b}{a}\right) \exp\left(-\frac{\left(\frac{x-b}{a}\right)^2}{2}\right), \quad a, b \in R, \quad a > 0.$$

2. МНАТ-вейвлет («Мексиканская шляпа») второго порядка

$$\psi(x) = (1-x^2) \exp\left(-\frac{x^2}{2}\right),$$

$$\psi_{ab}(x) = \frac{1}{\sqrt{a}} \left(1 - \left(\frac{x-b}{a}\right)^2\right) \exp\left(-\frac{\left(\frac{x-b}{a}\right)^2}{2}\right), \quad a, b \in R, \quad a > 0.$$

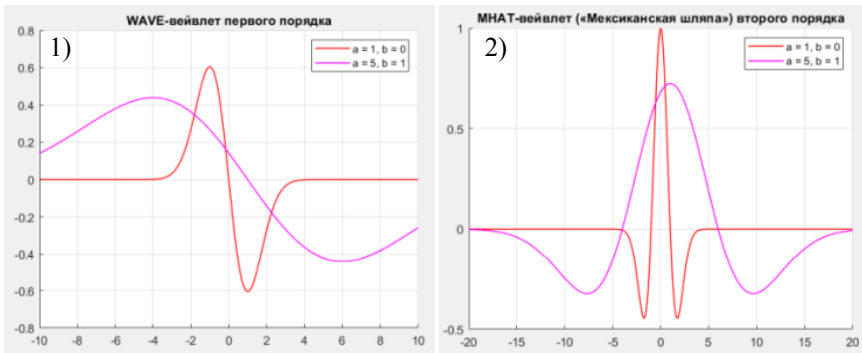


Рис. 1. Вейвлеты: 1) WAVE $\psi_{ab}(x)$, 2) МНАТ $\psi_{ab}(x)$ для $a = 1, b = 0$; $a = 5, b = 1$

На рис. 2 представлены результаты расчетов вейвлет-спектров для двух сигналов с использованием вейвлета второго порядка МНАТ, где

$$f(x) = \begin{cases} U, & \text{if } x_0 \leq x \leq x_0 + \delta \\ 0, & \text{иначе} \end{cases}, \quad a = 1..50, \quad b = 0..100,$$

$$\psi_{ab}(x) = \frac{1}{\sqrt{a}} \text{МНАТ}\left(\frac{x-b}{a}\right), \quad W(a, b) = \int_{-N}^N \psi_{ab}(x) f(x) dx, \quad N = 256.$$

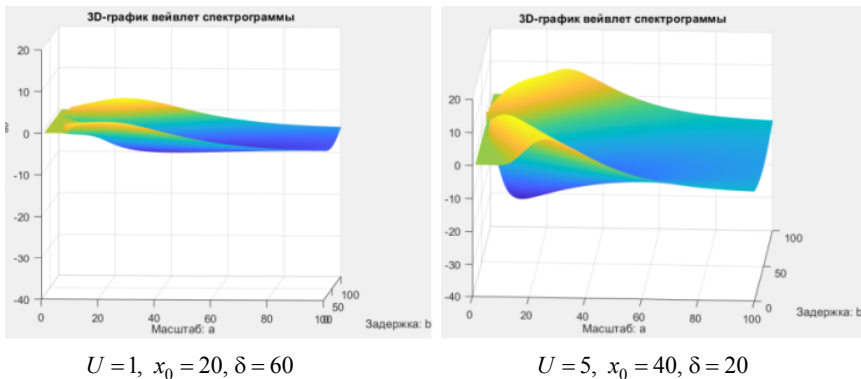


Рис. 2. Вейвлет-спектры сигналов

Синий цвет на рисунке соответствует минимальному значению вейвлет-коэффициентов $W(a, b)$, желтый – максимальному. Промежуточные значения соответствуют оттенкам зеленого и голубого различной интенсивности. Как следует из рисунков, при изменении параметров сигнала x_0 и δ меняются точки скачков амплитуды. Таким образом, полученные спектры позволяют проанализировать изменения амплитуды ВП для разных масштабов и сдвигов по времени.

Для построения сплайн-вейвлетов используются кардинальные сплайны $K_m^i(x)$, построенные с помощью B -сплайнов [7], имеющих компактные носители на целочисленной равномерной сетке $Z = \{\dots, -2, -1, 0, 1, 2, \dots\}$ с шагом $h = 1$:

$$K_m^i(x) = \frac{x-i}{m-1} K_{m-1}^i(x) + \frac{i+m-x}{m-1} K_{m-1}^i(x-1), \quad m > 1, \quad i \in \mathbb{R}.$$

$$K_1^i(x) = \begin{cases} 1, & i \leq x < i+1, \\ 0, & x \notin [i, i+1), \end{cases} \quad K_m^i(x-1) = \begin{cases} 1, & i+1 \leq x < i+2, \\ 0, & x \notin [i+1, i+2), \end{cases}$$

$$K_m^0(x) = K_m(x), \quad K_m(x) = \frac{x}{m-1} K_{m-1}(x) + \frac{m-x}{m-1} K_{m-1}(x-1), \quad m > 1,$$

$$K_1(x) = \varphi(x) = \begin{cases} 1, & 0 \leq x < 1 - \text{функция Хаара,} \\ 0, & \text{в остальных случаях.} \end{cases}$$

Точные значения функций $K_m(x)$, $m = \overline{2,5}$, на носителях с шагом $l = 1/2$ (для коэффициента масштабирования $N = 2$) приведены в табл. 1.

Таблица 1

x	0	$\frac{1}{2}$	1	3	2	$\frac{5}{2}$	3	$\frac{7}{2}$	4	$\frac{9}{2}$	5
$K_2(x)$	0	$\frac{1}{2}$	1	$\frac{1}{2}$	0	0	0	0	0	0	0
$K_3(x)$	0	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{1}{2}$	$\frac{1}{8}$	0	0	0	0	0
$K_4(x)$	0	$\frac{1}{48}$	$\frac{1}{6}$	$\frac{23}{48}$	$\frac{2}{3}$	$\frac{23}{48}$	$\frac{1}{6}$	$\frac{1}{48}$	0	0	0
$K_5(x)$	0	$\frac{1}{384}$	$\frac{1}{24}$	$\frac{76}{384}$	$\frac{11}{24}$	$\frac{230}{384}$	$\frac{11}{24}$	$\frac{76}{384}$	$\frac{1}{24}$	$\frac{1}{384}$	0

Точные значения функций $K_m(x)$, $m = \overline{2,4}$, на носителях с шагом $l = 1/3$ (для коэффициента масштабирования $N = 3$) приведены в табл. 2.

Таблица 2

x	0	$\frac{1}{3}$	$\frac{2}{3}$	1	$\frac{4}{3}$	$\frac{5}{3}$	2	$\frac{7}{3}$	$\frac{8}{3}$	3	$\frac{10}{3}$	$\frac{11}{3}$	4
K_2	0	$\frac{1}{3}$	$\frac{2}{3}$	1	$\frac{2}{3}$	$\frac{1}{3}$	0	0	0	0	0	0	0
K_3	0	$\frac{1}{18}$	$\frac{4}{18}$	$\frac{1}{2}$	$\frac{13}{18}$	$\frac{13}{18}$	$\frac{1}{2}$	$\frac{4}{18}$	$\frac{1}{18}$	0	0	0	0
K_4	0	$\frac{1}{162}$	$\frac{8}{162}$	$\frac{1}{6}$	$\frac{20}{54}$	$\frac{31}{54}$	$\frac{2}{3}$	$\frac{31}{54}$	$\frac{20}{54}$	$\frac{1}{6}$	$\frac{8}{162}$	$\frac{1}{162}$	0

Анализ сигнала на множестве вложенных линейных пространств $V_j \in L_2(R)$, таких, что

$$\dots \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \dots \subset V_j \subset \dots,$$

$$\bigcap_{j=-\infty}^{\infty} V_j = \{0\} \text{ и } \overline{\bigcup_{j=-\infty}^{\infty} V_j} = L_2(R), \quad f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1},$$

с ортонормированной системой масштабирующих функций и вейвлетов Хаара:

$$\begin{aligned}\varphi_{j,n}(x) &= \sqrt{2^j} \varphi(2^j x - n), \\ \text{supp}(\varphi_{j,n}) &= \left[\frac{n}{2^j}, \frac{(n+1)}{2^j} \right], \quad n = 0, 1, \dots, 2^j - 1; \\ \varphi_{j,n}(x) &= \sqrt{2^j} \psi(2^j x - n), \quad \text{supp}(\varphi_{j,n}) = \left[\frac{n}{2^j}, \frac{(n+1)}{2^j} \right], \\ n &= 0, 1, \dots, 2^j - 1, \quad \psi(x) = \varphi(2x) - \varphi(2x - 1),\end{aligned}$$

называется ортогональным кратномасштабным анализом (ОКА). Точность его представления возрастает с ростом уровня разрешения j . Отметим, что с ростом уровня разрешения j носители масштабирующих функций $\varphi_{j,n}(x)$ и вейвлетов Хаара $\psi_{j,n}(x)$ являются малыми порядка 2^{-j} .

Так как $V_{j+1} = V_j \oplus W_j$, где W_j – ортогональное дополнение к пространству V_j до V_{j+1} , то оператор проектирования $P_j : L_2(R) \rightarrow V_j$ на j -м уровне разрешения ($j \leq L$) представляется в виде суммы

$$\begin{aligned}P_j(f) &= \sum_{k \in \mathbb{Z}} a_{j-1,k} \varphi_{j-1,k}(x) + \sum_{k \in \mathbb{Z}} d_{j-1,k} \psi_{j-1,k}(x), \\ a_{j-1,k} &= (f, \varphi_{j-1,k}) = \int_R f(x) \overline{\varphi_{j-1,k}(x)} dx, \\ d_{j-1,k} &= (f, \psi_{j-1,k}) = \int_R f(x) \overline{\psi_{j-1,k}(x)} dx,\end{aligned}\tag{1}$$

где $a_{j-1,k}$ – коэффициенты аппроксимации $(j-1)$ -го уровня приближения; $d_{j-1,k}$ – детализирующие коэффициенты.

Точность представления сигнала при переходе на более низкие уровни декомпозиции (разрешения) снижается, но появляется возможность вейвлет-фильтрации и сжатия сигнала. Для произвольного j нормированные коэффициенты в (1) вычисляются по формулам [8]

$$\begin{aligned}a_{j-1,k} &= \frac{1}{\sqrt{2}} (a_{j,2k} + a_{j,2k+1}), \quad d_{j-1,k} = \frac{1}{\sqrt{2}} (d_{j,2k} + d_{j,2k+1}), \\ k &= 0, 1, \dots, 2^{j-1}.\end{aligned}\tag{2}$$

Соответственно для коэффициентов обратного преобразования получим

$$a_{j,2k} = \frac{1}{\sqrt{2}}(a_{j-1,k} + d_{j-1,k}), \quad a_{j,2k+1} = \frac{1}{\sqrt{2}}(a_{j-1,k} - d_{j-1,k}),$$

$$k = 0, 1, \dots, 2^{j-1}. \quad (3)$$

Максимально возможная глубина вейвлет-разложения L определяется количеством отсчетов сигнала $N = 2^L : L = \log_2 N$.

Выполним прямое и обратное вейвлет-разложение Хаара для функции

$$\varphi(x) = 4 \left[\frac{\sin\left(2\pi\left(x + \frac{3}{2}\right)\right) - \sin\left(\pi\left(x + \frac{3}{2}\right)\right)}{\pi\left(x + \frac{3}{2}\right)} \right] + \varepsilon \cdot \cos(20\pi x),$$

$$x \in [0, 1]; \quad N = 8, \quad \varepsilon \geq 0, \quad h = 1/N, \quad L = 3. \quad (4)$$

Для расчета коэффициентов прямого разложения используются формулы из табл. 3.

Таблица 3

j, i	Расчетные формулы прямого преобразования
3,0	$a_{3,k} = y_k, \quad d_{3,j} = 0, \quad k = \overline{1,8}$
2,1	$a_{21} = \frac{y_1 + y_2}{\sqrt{2}}, \quad d_{21} = \frac{y_1 - y_2}{\sqrt{2}}, \quad a_{22} = \frac{y_3 + y_4}{\sqrt{2}},$ $d_{22} = \frac{y_3 - y_4}{\sqrt{2}}, \quad a_{23} = \frac{y_5 + y_6}{\sqrt{2}}, \quad d_{23} = \frac{y_5 - y_6}{\sqrt{2}},$ $a_{24} = \frac{y_7 + y_8}{\sqrt{2}}, \quad d_{24} = \frac{y_7 - y_8}{\sqrt{2}}$
1,2	$a_{11} = \frac{a_{21} + a_{22}}{\sqrt{2}}, \quad d_{11} = \frac{a_{21} - a_{22}}{\sqrt{2}},$ $a_{12} = \frac{a_{23} + a_{24}}{\sqrt{2}}, \quad d_{12} = \frac{a_{23} - a_{24}}{\sqrt{2}}$
0,3	$a_{00} = \frac{a_{11} + a_{12}}{\sqrt{2}}, \quad d_{00} = \frac{a_{11} - a_{12}}{\sqrt{2}}$
	j – уровень разрешения, i – глубина разложения

Тогда получим нормированный вектор сигнала (4):

$$\left[y_1 = 0.18, y_2 = -0.008, y_3 = -0.078, y_4 = -0.039, \right. \\ \left. y_5 = 0.034, y_6 = 0.064, y_7 = 0.009, y_8 = -0.111 \right].$$

Исходный сигнал, аппроксимирующая и детализирующая функции для различных глубин разложения ($i = \overline{0,3}$) представлены на рис. 3.

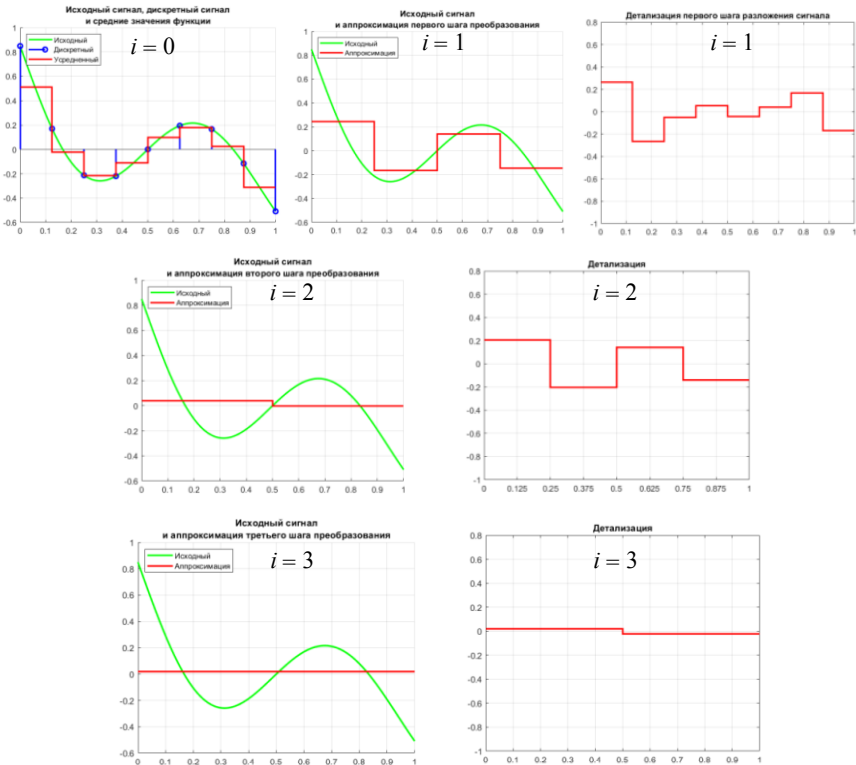


Рис. 3. Аппроксимирующие и детализирующие функции сигнала (4) для различных глубин разложения

Отметим, что при сложении аппроксимирующей функции с детализирующей на любом этапе получается аппроксимирующая функция предыдущего этапа. Восстановление сигнала осуществляется по фор-

мулам обратного преобразования (3). С увеличением количества отсчетов N возрастает количество этапов представления аппроксимирующей и детализирующей части КМА. Например, на рис. 4 для рассматриваемой функции $\varphi(x)$, $N = 32$, $\varepsilon = 0$ представлены детализирующие функции для разных глубин разложения i .

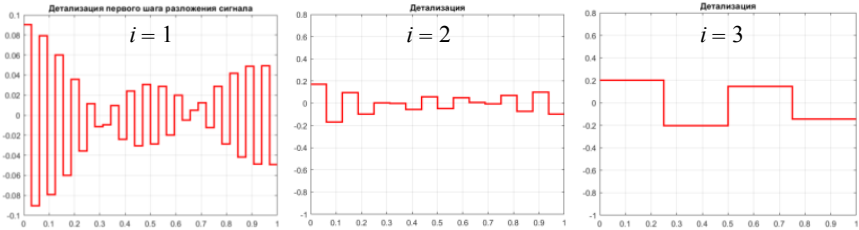


Рис. 4. Детализирующие функции $\varphi(x)$, $N = 32$, $\varepsilon = 0$.

КМА позволяет удалять высокочастотные составляющие (шумы) из спектра сигнала, вводя, например, пороговую функцию [9,10]

$$d = \begin{cases} d_{jk}, & \text{если } |d_{jk}| \geq \tau, \\ 0, & \text{если } |d_{jk}| < \tau. \end{cases}$$

На рис. 5 представлены результаты фильтрации $\varphi(x, \varepsilon)$ для $\varepsilon = 0.05$ и различных пороговых значений τ при $N = 32$, α – максимальное отклонение $\varphi(x, \varepsilon)$ от $\varphi(x, 0)$.

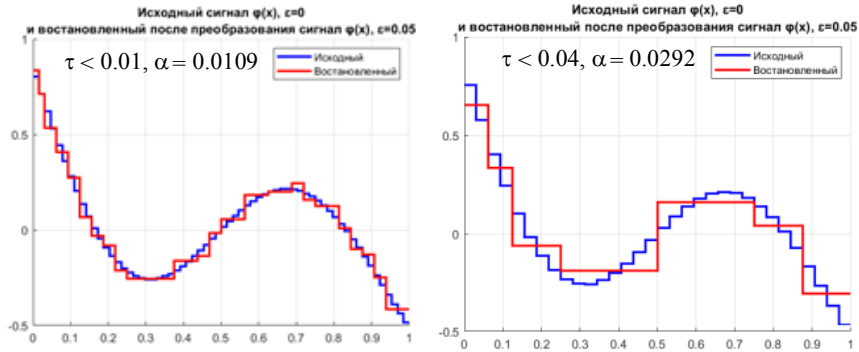


Рис. 5. Пороговое сглаживание, α – максимальное отклонение

На рис. 6 представлены результаты фильтрации сигнала $\varphi(x, \varepsilon)$, $\varepsilon = 0.25$ для различных значений τ для $N = 64$.

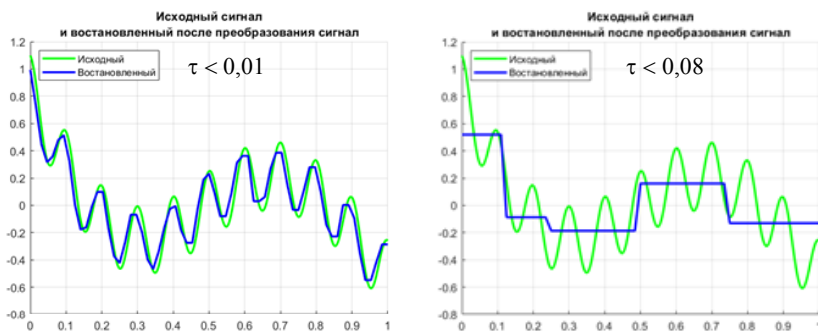


Рис. 6. Пороговое сглаживание для $\varphi(x)$, $\varepsilon = 0.25$

С увеличением параметра τ эффективность сглаживания уменьшается. Вычислительная сложность алгоритма КМА: $\mu = O(NL)$, $N = 2^L$.

ЛИТЕРАТУРА

1. Grossman A., Morlet J. Decomposition of Hardy functions into square integrable wavelets of constant shape // SIAM J. Math. – 1984. – P. 723–736.
2. Чу Ч. Введение в вейвлеты. – М.: Мир, 2001. – 412 с.
3. Добеши И. Десять лекций по вейвлетам. – М.: РХД, 2001.
4. Zadeh Lotfi A. Fuzzy Logic, Neural Networks, and Soft Computing // Communications of the ACM. – 1994. – V. 37. – No. 3. – P. 77–84.
5. Смоленцев Н.К. Основы теории вейвлетов. Вейвлеты в MATLAB. – М.: ДМК Пресс, 2014. – 628 с.
6. Нагорнов О.В. и др. Вейвлет-анализ в примерах: учеб. пособие. – М.: НИЯУ МИФИ, 2010. – 120 с.
7. Де Бор К. Практическое руководство по сплайнам: пер. с англ. – М.: Радио и связь, 1985. – 304 с.
8. Дремин И.М., Иванов О.В., Нечитайло В.А. Вейвлеты и их использование // УФН. – 2001. – Т. 171. – С. 465–501.
9. Меркушева А.В. Фильтрация нестационарного сигнала (речи) в вейвлет-области с адаптацией к виду и динамике шума // Научное приборостроение. – 2003. – Т. 13. – № 2. – С. 73–87.
10. Воскобойников Ю.Е. Вейвлет-фильтрации сигналов и изображений (с примерами в пакете MathCAD). – Новосибирск: НГАСУ (Сибстрин), 2015. – 188 с.
11. Малла С. Вейвлеты в обработке сигналов. – М.: Мир, 2005. – 671 с.

Берцун Владимир Николаевич, канд. физ.-мат. наук, доцент кафедры вычислительной математики и компьютерного моделирования ММФ НИ ТГУ; bvn@math.tsu.ru.

Научное издание

**ДЕСЯТАЯ СИБИРСКАЯ КОНФЕРЕНЦИЯ
ПО ПАРАЛЛЕЛЬНЫМ И
ВЫСОКОПРОИЗВОДИТЕЛЬНЫМ
ВЫЧИСЛЕНИЯМ**

Сборник статей

(Томск, 5–7 октября 2021 г.)

Редактор *Л.А. Пивоварова*

Дизайн, верстка *Д.В. Фортеса*

ООО «Издательство научно-технической литературы»
634034, Томск, ул. Студенческая, 4, тел. (3822) 53-10-35

Изд. лиц. ИД № 04000 от 12.02.2001. Подписано к печати 07.12.2021.
Формат 60 × 84 ¹/₁₆. Бумага офсетная. Печать офсетная. Гарнитура «Таймс».
Усл. п. л. 7,09. Уч.-изд. л. 7,94. Тираж 50 экз. Заказ № 23.
