

Параллельные алгоритмы для решения обратных задач переноса примеси

А.В. Старченко, Е.А. Панасенко

При решении обратных задач переноса примеси используются сопряженные уравнения и двойственное представление функционала концентрации примеси. Параллельная реализация строится на основе следующих подходов: геометрическая декомпозиция, функциональная декомпозиция, комбинация функциональной и геометрической декомпозиций.

1. Введение

В настоящее время проблема охраны окружающей среды и ее восстановления становится одной из важных задач науки, развитие которой стимулируется всевозрастающими темпами технического прогресса во всех странах мира. Ухудшение качества атмосферного воздуха, накопление в нём газообразных компонентов происходит вследствие выбросов в атмосферу отходов промышленных предприятий и выхлопных газов автотранспорта, что приводит к ухудшению здоровья населения, а в глобальном масштабе - к изменению климата на планете. При этом наиболее значительный вклад в загрязнение воздуха вносит автотранспорт. Особенность автотранспорта заключается в том, что он является подвижным источником загрязнения, что проявляется в низком его расположении и непосредственной близости к зонам жилой застройки. Все это приводит к тому, что автотранспорт создает в городах обширные зоны, в пределах которых предельно-допустимая концентрация загрязняющих субстанций в атмосферном воздухе превышена в несколько раз.

Поступление загрязняющих веществ в атмосферу избежать невозможно, но разумное использование природных ресурсов и постоянный контроль качества атмосферного воздуха позволяют обеспечить безопасный уровень воздействия на атмосферу и избежать глобально негативных последствий.

Одним из способов оценки и прогноза уровня загрязнения воздуха является контроль интенсивности выбросов вредных веществ, который производится с помощью постов наземных наблюдений. Но даже разветвленная сеть таких пунктов наблюдения не всегда может предоставить достоверную информацию для природоохранных служб. Большую помощь здесь может оказать применение методов математического моделирования и, особенно, технологии численного решения обратных задач по определению характеристик источников загрязнения атмосферного воздуха по данным измерений концентрации вредных веществ. Но сложность и взаимосвязанность процессов распространения и переноса загрязняющих субстанций, происходящих в атмосферном воздухе, делают модели оценки и прогнозирования качества воздуха громоздкими в математической записи и весьма требовательными к вычислительным ресурсам. Перспективным способом решения этих проблем является использование современных высокопроизводительных многопроцессорных вычислительных систем, которые обеспечивают существенное ускорение получения результатов расчетов и повышение качества численного прогноза.

Целью данной работы является создание новых вычислительных алгоритмов для решения обратных задач переноса примеси на суперкомпьютерах с параллельной архитектурой.

2. Физическая постановка основной задачи переноса примеси

Распространение примеси в приземном слое атмосферы над промышленным центром и его окрестностями исследуется при следующих условиях. Рассматривается ограниченный в прямоугольнике участок территории, на котором в течение времени исследования происходит выброс от различных источников. Мощность источников зависит от времени. Распределение по облас-

ти исследования поступающей примеси определяется метеорологическими условиями. Кроме того, сделано предположение, что примесь не вступает в химические реакции с другими веществами и скорость ее переноса совпадает со скоростью движения окружающего воздуха (рис. 1).

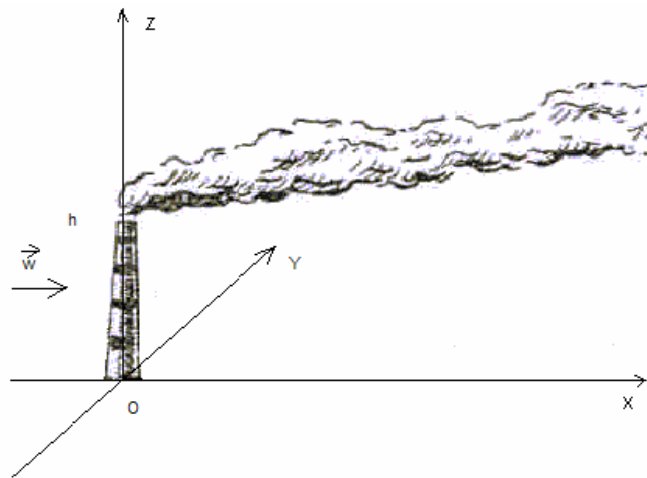


Рис. 1. Распределение поступающей примеси по области исследования

3. Математическая постановка основной задачи переноса примеси

С учетом принятой физической постановки задачи, адвективно-диффузионное уравнение, моделирующее перенос газообразной примеси в заданном потоке, представляется в следующем виде:

$$\frac{\partial C}{\partial t} + U \frac{\partial C}{\partial x} + V \frac{\partial C}{\partial y} + W \frac{\partial C}{\partial z} + \alpha C = \frac{\partial}{\partial x} \left[\Gamma \frac{\partial C}{\partial x} \right] + \frac{\partial}{\partial y} \left[\Gamma \frac{\partial C}{\partial y} \right] + \frac{\partial}{\partial z} \left[K_z \frac{\partial C}{\partial z} \right] + Q, \quad (1)$$

где C - концентрация примеси; U, V, W - компоненты соленоидального вектора скорости атмосферного воздуха; Γ, K_z - коэффициенты турбулентной диффузии; Q - интенсивность поступления примеси от источников.

Начальные и граничные условия имеют следующий вид:

$$\begin{aligned} t = 0: C(0, x, y, z) &= C_0(x, y, z); \\ x = 0: \frac{\partial C}{\partial x} &= 0; \quad x = L_x: \frac{\partial C}{\partial x} = 0; \\ y = 0: \frac{\partial C}{\partial y} &= 0; \quad y = L_y: \frac{\partial C}{\partial y} = 0; \\ z = 0: K_z \frac{\partial C}{\partial z} &= \alpha C - R; \quad z = L_z: \frac{\partial C}{\partial z} = 0. \end{aligned} \quad (2)$$

4. Физическая постановка обратной задачи переноса примеси

Требуется по известным метеорологическим параметрам атмосферы и результатам измерений концентрации примеси в N точках, проводимых в течение некоторого периода времени T , определить параметры (мощность, координаты и время срабатывания) источников примеси (рис. 2).

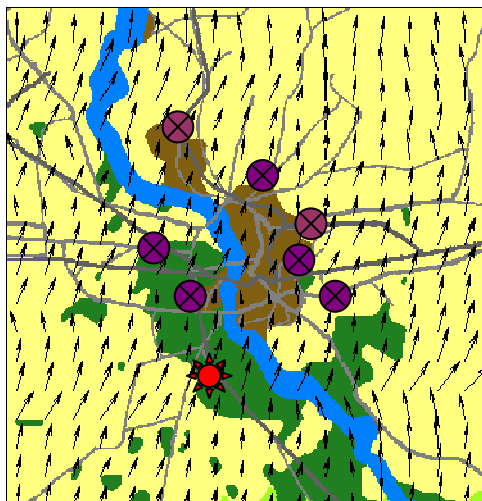


Рис. 2. Определение местонахождения источника выброса

5. Математическая постановка обратной задачи

В данной работе для решения обратных задач привлекается подход Г.И.Марчука [1], который основан на решении уравнения, сопряженного с полуэмпирическим уравнением турбулентной диффузии (1), и двойственным представлением функционала от концентрации примеси.

Математическая постановка обратной задачи переноса примеси включает несколько одно-типных сопряженных уравнений (их количество совпадает с количеством измерений концентрации) и поиск минимума некоторого функционала в процессе решения сопряженных уравнений.

Сопряженная к (1) и (2) задача получается следующим образом: уравнение (1) умножается на функции $C_k^* = C_k^*(t, x, y, z)$ и интегрируется по времени и пространству. Тогда используя интегрирование по частям, приходим к сопряженным постановкам:

$$-\frac{\partial C_k^*}{\partial t} - \frac{\partial UC_k^*}{\partial x} - \frac{\partial VC_k^*}{\partial y} - \frac{\partial WC_k^*}{\partial z} + \sigma C_k^* - \frac{\partial}{\partial x} \left[\Gamma \frac{\partial C_k^*}{\partial x} \right] - \frac{\partial}{\partial y} \left[\Gamma \frac{\partial C_k^*}{\partial y} \right] - \frac{\partial}{\partial z} \left[K_z \frac{\partial C_k^*}{\partial z} \right] = P_k; \quad (3)$$

$$k = 1, \dots, N;$$

с соответствующими начальными и граничными условиями:

$$C_k^*(T, x, y, z) = 0;$$

$$x = 0 : UC_k^* + \Gamma \frac{\partial C_k^*}{\partial x} = 0; \quad x = L_x : UC_k^* + \Gamma \frac{\partial C_k^*}{\partial x} = 0;$$

$$y = 0 : VC_k^* + \Gamma \frac{\partial C_k^*}{\partial y} = 0; \quad y = L_y : VC_k^* + \Gamma \frac{\partial C_k^*}{\partial y} = 0; \quad (4)$$

$$z = 0 : K_z \frac{\partial C_k^*}{\partial z} = \alpha C_k^*; \quad z = L_z : \frac{\partial C_k^*}{\partial z} = 0,$$

где $P_k = \delta(x - x_k)\delta(y - y_k)\delta(z - z_k)\delta(t - t_k)$, N – количество точек наблюдения с координатами (x_k, y_k, z_k) , t_k – момент времени измерения концентрации, и двойственному представлению функционала [2]:

$$J(C) = \int_0^T \int_0^{L_x} \int_0^{L_y} \int_0^{L_z} CP_k dz dy dx dt = C_k = \int_0^T \int_0^{L_x} \int_0^{L_y} \int_0^{L_z} C_k^* Q dz dy dx dt + \int_0^T \int_0^{L_x} \int_0^{L_y} C_k^* \Big|_{z=0} R dy dx dt; \quad k = 1, \dots, N. \quad (5)$$

При этом нахождение решения систем уравнений (3)-(4) и (5) позволит установить параметры (мощность, координаты и время срабатывания) источника загрязнения атмосферного воздуха.

6. Численное решение сопряженной задачи

Для численной реализации задачи (3) – (4) использовались метод конечного объема и явные разностные схемы, для аппроксимации адвективных членов применялась схема MLU Ван Лира [Ошибка! Источник ссылки не найден.].

При построении разностного аналога дифференциальной задачи сначала строилась сетка. Для этого выбирались шаги $h_x, h_y, h_z > 0$ и на прямоугольной области $L_x \times L_y \times L_z$ задавалась сетка:

$$\omega_{h_x, h_y, h_z} = \{x_i = ih_x, y_j = jh_y, z_l = lh_z\}, i = 0, 1, \dots, N_x, j = 0, 1, \dots, N_y, l = 0, 1, \dots, N_z\}.$$

Построенная сетка используется при определении векторов скорости и концентрации.

7. Параллельная реализация сопряженной задачи

При численном решении задач переноса примеси выделяют два основных способа параллельной реализации: распараллеливание по физическим процессам (функциональная декомпозиция) и геометрическая декомпозиция расчетной области. При первом способе реализации вычислительные процессы выполняются параллельно и при этом каждый из них занят выполнением принципиально различных задач.

При использовании же декомпозиции расчетной области область исследования делится на подобласти, число которых равно числу процессов, и каждый процесс ведет расчеты в своей подобласти независимо от других. При таком подходе объемы вычислений, которые осуществляет каждый процесс, очень близки между собой [Ошибка! Источник ссылки не найден.].

При решении обратной задачи определения характеристик источников на каждом шаге по времени решается не одна, а N независимых сопряженных задач, следовательно, они могут решаться параллельно.

Такие условия проведения численного моделирования позволяют привлекать высокопроизводительную вычислительную технику, в частности, вычислительный кластер ТГУ SKIF Cyberia.

7.1 Функциональная декомпозиция

Распараллеливание по физическим процессам метода численного решения задачи по определению параметров точечных мгновенных источников производилось с использованием принципа «master-slave» (рис. 3). В этом случае управляющий master-процесс с частотой ω передает каждому slave-процессу значения метеорологических параметров, необходимые для решения сопряженных задач. Подчиненные slave-процессы, получив данные, в свою очередь, ведут расчеты независимо друг от друга, и найденные на каждом шаге по времени приближенные решения своей сопряженной задачи (3) – (4) возвращают управляющему процессу, который ищет глобальный минимум функционала

$$\Phi(C) = \sum_{k=1}^N \left(\int_0^T \int_0^{L_x} \int_0^{L_y} \int_0^{L_z} C_k^* Q dz dy dx dt + \int_0^T \int_0^{L_x} \int_0^{L_y} C_k^* \Big|_{z=0} R dy dx dt - C_k \right)^2, \quad (6)$$

что в итоге позволит определить параметры источников загрязнения.

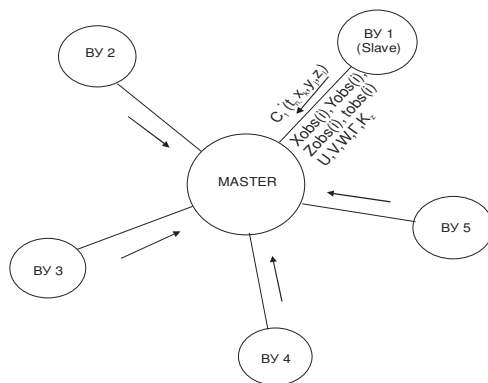


Рис. 3. Схема параллельной реализации

Заметим, что при таком способе организации параллельных вычислений при запуске параллельной программы на $p+1$ процессах большую долю вычислительной работы обычно выполняют p подчиненных slave-процессов. Один управляющий master-процесс координирует работу остальных (подготавливает и рассылает slave-процессам данные для расчета, собирает результаты их расчетов и осуществляет дополнительную обработку этих результатов), причем каждый slave-процесс в данный текущий момент целиком занят выполнением одной задачи.

В работе на основе теоретических оценок и в результате вычислительных экспериментов исследовано ускорение и эффективность данного параллельного алгоритма для задачи идентификации параметров мгновенного источника. Теоретические выкладки приводят к следующим результатам:

$$S_p^{\Phi 1} \approx \frac{p-1}{1 + \frac{5\chi \cdot \omega}{m}}, E_p^{\Phi 1} = \frac{S_p}{p} \approx \frac{1-1/p}{1 + \frac{5\chi \cdot \omega}{m}}. \quad (7)$$

Здесь p - число используемых процессов, m - число арифметических операций для вычисления одного значения сеточной функции, χ - отношение времени межпроцессорной передачи одного числа к времени арифметической операции, $\Phi 1$ - первый вариант функциональной декомпозиции.

Из (7) видно, что при таком способе организации параллельных вычислений не удастся достичь идеального параллелизма, поскольку нет хорошей балансировки загрузки используемых в расчетах p процессоров. Это связано с тем, что master-процесс не участвует в проведении массовых вычислений (не решает отдельную сопряженную задачу). Поэтому с целью повышения эффективности алгоритма рассматривается другой подход, в котором master-процесс дополнительно к своей работе также решает одну из N сопряженных задач (3) – (4). В этом случае теоретическая оценка ускорения и эффективности может быть записана следующим образом:

$$S_p^{\Phi 2} \approx \frac{p}{1 + \frac{5\chi \cdot \omega}{m}}, E_p^{\Phi 2} \approx \frac{1}{1 + \frac{5\chi \cdot \omega}{m}} \quad (8)$$

и разработанная параллельная программа будет иметь более высокие показатели по времени выполнения ($\Phi 2$ - второй вариант функциональной декомпозиции).

Для подтверждения полученных оценок ускорения параллельного алгоритма был проведен ряд расчетов на кластере ТГУ СКИФ Cyberia. Результаты представлены в таблице 1.

Таблица 1. Время счета (сек) параллельной программы для рассматриваемых вариантов функциональной декомпозиции

Количество измерений, N	Число используемых процессов/время счета для варианта $\Phi 1$	Число используемых процессов/время счета для варианта $\Phi 2$
---------------------------	--	--

5	6/312	5/323
10	11/316	10/337
20	21/322	20/354

Из таблицы видно, что оба рассматриваемых варианта имеют высокие показатели масштабируемости, что выражается в слабом изменении времени счета при добавлении процессорных элементов в связи с увеличением объема вычислений (N). Однако первый вариант функциональной декомпозиции алгоритма решения обратной задачи (3)-(5) имеет небольшое преимущество во времени выполнения параллельной программы, что обусловлено появлением дополнительных возможностей совмещения работы активных процессов, количество которых для одних и тех же условий на единицу больше, чем во втором варианте. В то же время соотношение эффективности рассматриваемых параллельных программ $e^\Phi = \frac{E_p^{\Phi 1}}{E_p^{\Phi 2}} < 1$ показывает лучшую масштабируемость второго варианта декомпозиции. Что хорошо согласуется с теоретическими оценками (7)-(8), согласно которым $(e^\Phi)_{теор} \approx \frac{p}{p+1}$.

7.2 Геометрическая декомпозиция

При решении каждой сопряженной задачи (3)-(4) в данной работе также использовалась одномерная (вдоль оси OY , рис. 4) геометрическая декомпозиция сеточной области. Этот способ распараллеливания численного алгоритма решения нестационарного адвективно-диффузионного уравнения показал высокую эффективность.

Рассмотрим параллельную реализацию алгоритма решения обратной задачи переноса смеси при использовании принципа геометрической декомпозиции сеточной области. В этом случае вычисления осуществляются следующим образом: N сопряженных задач (3) – (4) последовательно одна за другой решаются численно в подобласти вычислительной сетки, распределенной каждому процессу. Вычисления в течение промежутка времени $[0, T]$ проводятся каждым процессом одновременно, но для корректного расчета сеточных значений функции C_k^* вблизи границ подобласти (рис. 4) требуются некоторые сеточные значения с соседних по расположению подобластей. Это может быть обеспечено лишь путём межпроцессорной передачи данных с использованием, например, стандарта MPI [Ошибка! Источник ссылки не найден., Ошибка! Источник ссылки не найден.].

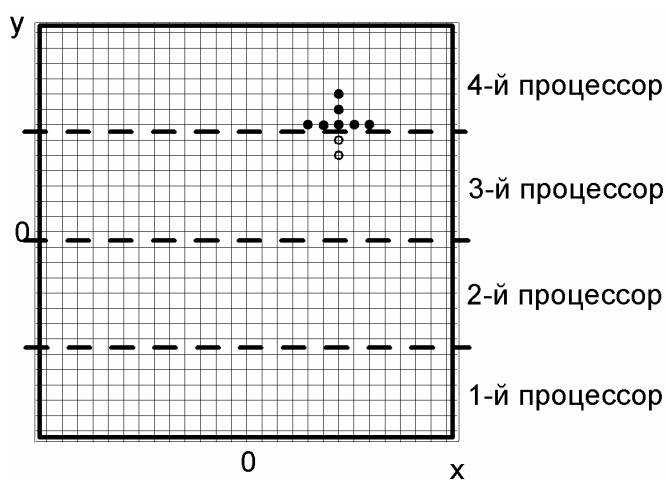


Рис. 4. Расчетная сетка в плоскости xOy с указанием распределения подобластей по четырем процессорам. На приграничной сеточной линии приведен разностный шаблон. Открытые кружки указывают на сеточные значения, которые необходимо получить с соседнего процессора

Теоретическая оценка ускорения и эффективности данного алгоритма имеет следующий вид:

$$S_p^{\Gamma} \approx \frac{p}{1 + \frac{\chi}{m} \left(\frac{5\omega}{N} + \frac{4p}{N_x} \right)}, E_p^{\Gamma} = \frac{S_p^{\Gamma}}{p} \approx \frac{1}{1 + \frac{\chi}{m} \left(\frac{5\omega}{N} + \frac{4p}{N_x} \right)}, \quad (9)$$

где N_x - размер вычислительной сетки по оси Ox , N - количество решаемых сопряженных задач.

Поскольку в полученной оценке эффективности рассматриваемого способа распараллеливания вклад межпроцессорной передачи данных во временные затраты более значительный, стоит ожидать, что применение геометрической декомпозиции при решении обратных задач переноса примеси, в которых на каждом шаге по времени требуется передача данных master – процессу, менее перспективно, чем использование функциональной декомпозиции.

Для подтверждения сделанных выводов были проведены вычислительные эксперименты, результаты которых представлены в таблице 2.

Таблица 2. Время счета параллельной программы для геометрической декомпозиции

Количество измерений, N	Число используемых процессов, p	Время, сек
5	5	627
10	10	643
20	20	743

Сопоставляя результаты, приведенные в таблицах 1 и 2 видно, что применение геометрической декомпозиции для рассматриваемой задачи идентификации параметров внезапного выброса в атмосферу по данным измерений существенно (более чем в два раза) уступает по быстродействию функциональной декомпозиции. Причиной этого являются коммуникационные затраты на межпроцессорную передачу данных, необходимую для корректного решения сопряженных задач (3) – (4) на всей сеточной области, разделенной между активными процессами [Ошибка! Источник ссылки не найден.].

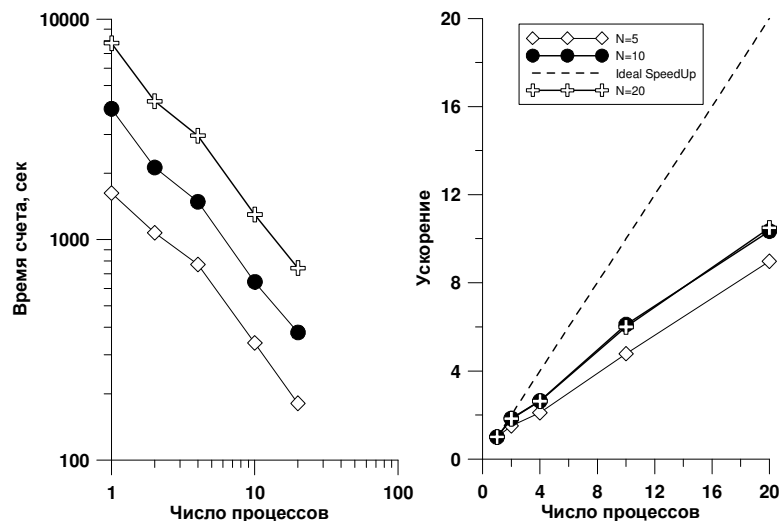


Рис. 5. Время счета (слева) и ускорение (справа) параллельной программы решения обратной задачи переноса примеси, базирующейся на принципе одномерной декомпозиции расчетной области

Данные, представленные на рисунке 5, показывают, что при увеличении числа активных процессов время счета уменьшается, причем пропорционально, для различного N . В целом, этот способ распараллеливания алгоритма решения обратной задачи для рассмотренных здесь

условий обеспечивает 50% эффективность параллельной программы для числа процессов $p = 2, 4, 10, 20$, что свидетельствует о неплохой масштабируемости параллельной реализации для $p \leq 20$.

7.3 Комбинированный способ распараллеливания

Рассмотренные выше подходы создания параллельных версий алгоритмов решения обратной задачи переноса примеси показали неплохие результаты по эффективности, однако количество используемых активных процессов в них ограничено: в случае функциональной декомпозиции число проведенных измерений N , а при одномерной геометрической декомпозиции размером сетки N_x и выбранным сеточным шаблоном.

При использовании комбинированного подхода параллельной реализации предлагается для увеличения в расчетах количества используемых активных процессов совместить применение функциональной и геометрической декомпозиции. Все p активных процессов делятся на N групп (по числу измерений), каждая из которых решает численно одну сопряженную задачу (3),(4), используя одномерную геометрическую декомпозицию. Для получения приближенного решения процессы отдельной группы обмениваются между собой рассчитанными значениями сеточной функции вблизи границ подобластей. Поиск минимума функционала выполняет один из p процессов, которому остальные передают результаты промежуточных расчетов.

Теоретические оценки такого способа параллельной реализации записываются в следующем виде:

$$S_p^{комб} \approx \frac{p}{1 + \frac{\chi}{m} \left(5\omega + \frac{4p}{N \cdot N_x} \right)}, E_p^{комб} = \frac{S_p^{комб}}{p} \approx \frac{1}{1 + \frac{\chi}{m} \left(5\omega + \frac{4p}{N \cdot N_x} \right)}, \frac{p}{N} = 1, 2, 3, \dots \quad (10)$$

Сравнивая оценки параллельных реализаций (9) и (10) можно отметить, что эффективность комбинированного подхода не хуже, чем при геометрической декомпозиции, однако количество используемых активных процессов можно увеличить на порядок, что при имеющей место масштабируемости параллельной версии алгоритма должно привести к значительному сокращению времени счета.

Результаты вычислительных экспериментов для задачи идентификации характеристик мгновенного источника подтвердили предварительные теоретические оценки (таблица 3).

Таблица 3. Время счета параллельной программы для комбинации функциональной и геометрической декомпозиции

Количество измерений, N	Время счета последовательной программы, сек	Число используемых процессов/Время счета для комбинированного подхода, сек
5	1626	100/33
10	3922	200/42
20	7799	400/59

Стоит отметить, что результаты численных расчетов при использовании комбинированного подхода показывают снижение времени счета с уменьшением вычислительной трудоемкости задачи, замедление ускорения параллельной программы с ростом числа используемых процессов и при сокращении количества сопряженных задач N . Комбинированный подход действительно на порядок позволяет увеличить число используемых процессов, при этом эффективность параллельной программы остается на уровне не менее 50%, что является неплохим для задач такого класса показателем. Однако, в целом, функциональная декомпозиция имеет существенное преимущество, судя по (8), (9) и (10).

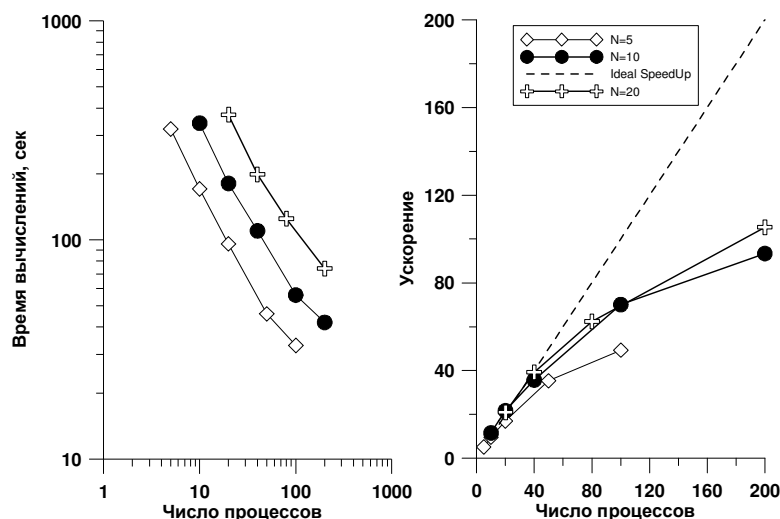


Рис. 6. Время счета (слева) и ускорение (справа) параллельной программы решения обратной задачи переноса примеси (комбинированный способ распараллеливания)

8. Заключение

В работе с единых позиций сформулированы математические постановки для решения обратных задач по определению параметров источника загрязнения атмосферного воздуха по измеренным значениям концентрации примеси. Математические постановки опираются на аппарат сопряженных уравнений и двойственное представление функционала от концентрации примеси. Для численного решения сопряженных задач используются конечно-разностные методы, явные разностные схемы, метод конечного объема. Для ускорения численного решения задачи обоснованы и применены следующие способы параллельной реализации задач переноса примеси: функциональная декомпозиция, геометрическая декомпозиция и комбинированный способ. Проведены оценки ускорения и эффективности всех трех подходов параллельной реализации. По произведенным оценкам ускорения и эффективности способов параллельной реализации получено, что, в целом, функциональная декомпозиция имеет существенное преимущество по сравнению с остальными подходами.

ЛИТЕРАТУРА

1. Марчук Г.И. Математическое моделирование в проблеме окружающей среды. М.: Наука, 1982.
2. Панасенко Е.А., Старченко А.В. Определение городских районов-загрязнителей атмосферного воздуха по данным наблюдений // Оптика атмосферы и океана. 2009. Т.22, № 03. С. 279-283.
3. Van Leer B. Towards the ultimate conservative difference scheme. II. Monotonicity and conservation combined in a second order scheme // J. Comput. Phys. 1974. V.14. P. 361 – 370.
4. Воеводин В.В., Воеводин Вл.В. Параллельные вычисления. СПб.: БХВ – Петербург, 2002.
5. Message Passing Interface Forum. MPI: A message – Passing Interface Standart // International Journal of Super Computer Applications. 1994. Vol. 8(314). P. 165 – 414.
6. Малышкин В.Э., Корнеев В.Д. Параллельное программирование мультикомпьютеров. Н-ск: изд-во НГТУ, 2006.
7. Старченко А.В. Параллельные вычисления в задачах охраны окружающей среды // Вторая Сибирская школа-семинар по параллельным вычислениям: Сб. науч. тр. / Томск: Изд-во Том. ун-та, 2004. С. 17-22.