

Метод поиска интервалов неопределенности кинетических констант химической реакции с использованием параллельных вычислительных систем

Э.Р. Ахматсафина, И.М. Губайдуллин, С.И. Спивак

При решении обратной задачи химической кинетики получаем многомерную область для кинетических констант, которая удовлетворяет рассматриваемой математической модели в пределах погрешности измерений, т. е. по каждой из констант имеем интервал неопределенности. В работе рассматривается задача поиска интервалов неопределенности для кинетических констант химической реакции циклоалюминирования олефинов с использованием параллельных вычислительных систем. Описан и обоснован метод распараллеливания для решения задачи поиска интервалов неопределенности. Приведены результаты расчетов.

1. Введение

Традиционно, алюминийорганические соединения широко применяются для осуществления таких реакций как термическое гидроалюминирование олефинов, диенов и ацетиленов, восстановление карбонильных соединений, сложных эфиров и нитрилов. В ряду этих исследований к числу наиболее крупных достижений отечественной и мировой науки последних лет в области химии алюминийорганических соединений следует отнести открытие реакции каталитического циклоалюминирования непредельных соединений (*Реакция Джемилева*) с помощью простейших триалкил- и алкилгалогеналанов с участием Ti- и Zr-содержащих комплексных катализаторов, позволяющей синтезировать ранее неописанные классы трех-, пяти- и макроциклических алуминакарбоциклов.

Изучение механизмов сложных реакций металлокомплексного катализа, к которым относятся реакции циклоалюминирования, на сегодняшний день представляет огромный практический интерес. При этом возникают как физико-химические, так и математические проблемы. Физико-химические проблемы сводятся к тому, что очень трудно, а иногда практически не представляется возможным на существующем оборудовании измерить характеристики промежуточных соединений реакции, из чего следует математическая неоднозначность решения обратных задач определения кинетических констант. Поэтому в решении обратной задачи получаем многомерную область для кинетических констант, которая удовлетворяет рассматриваемой модели в пределах погрешности измерений. Математические проблемы предполагают значительный объем вычислений при поиске этой многомерной области для констант. Для повышения эффективности решения данной задачи целесообразно использовать высокопроизводительные параллельные вычислительные системы.

Целью исследования является описание метода поиска интервалов неопределенности кинетических констант с использованием многопроцессорных технологий, поиск интервалов неопределенности констант скоростей для реакции циклоалюминирования олефинов, физико-химическая интерпретация полученных результатов. Для проведения расчетов использовался суперкомпьютер Уфимского государственного авиационно-технического университета.

2. Постановка задачи и схема поиска областей неопределенности

Классическая постановка задач математической обработки эксперимента восходит к работам Лежандра и Гаусса и состоит в поиске таких значений параметров математического описания, при которых минимизируется какой-либо критерий соответствия расчета измерениям, например, сумма квадратов отклонений экспериментально измеренных и рассчитанных по моде-

ли величин: $E = \sum_{j=1}^J \sum_{i=1}^I (x_{ij}^o - x_{ij}^p)^2$. Такой способ широко известен как метод наименьших квадратов.

Определение точки, соответствующей глобальному минимуму критерия – достаточно сложная вычислительная задача. Возникает вопрос: чем константы, соответствующие минимуму критерия, предпочтительнее констант, соответствующих некоторой окрестности критерия, если при этом сохраняется достаточная точность описания эксперимента, сопоставимая с его погрешностью? Этот вопрос об определении области в пространстве параметров, каждая точка внутри которого достаточно хорошо описывает измерения, был поставлен Л. В. Канторовичем [1] как решение системы неравенств:

$$\sum_{i=1}^I \sum_{j=1}^J |x_{ij}^s - x_{ij}^p| \leq \epsilon_i \quad (1)$$

Величины ϵ_i в системе неравенств (1) есть характеристики предельно допустимой погрешности эксперимента, информация о величине которой, как правило, присутствует. И тогда выполнение условий (1) означает, что модель описывает измерения в пределах, обусловленных величиной максимально допустимой погрешности измерений, что совершенно естественно.

Рассмотрение многомерной области для констант скоростей начнем с одномерного и двумерного случаев. Для одномерного случая область представляет собой интервал неопределенности. Определим интервал неопределенности как отрезок, вариация константы внутри которого сохраняет совместность системы неравенств (1). Для определения интервала по K -й константе необходимо решить две задачи математического программирования: найти $\min K$ и $\max K$ при выполнении ограничений (1).

В начальный момент времени нам известна начальная точка, принадлежащая искомому интервалу – точка, полученная в решении обратной задачи. Необходимо найти интервал по каждой из констант, беря начальную точку за исходную в поиске. Предлагается следующий алгоритм. Пусть ищется интервал неопределенности по K_i константе. Чтобы найти границы искомого интервала, нужно зафиксировать все константы, кроме K_i . Для этой константы выберем достаточно большой отрезок (например, $[0; 2K_i]$), содержащий исходную точку, покроем его равномерной сеткой с определенным достаточно малым шагом. В узлах сетки получим различные значения K_i константы, в каждом из узлов проверим выполнение системы неравенств (1). Если неравенства верны, то данный узел принадлежит искомому интервалу. Определяя наименьшее и наибольшее значения принадлежащих интервалу узлов, получаем значения границ интервала неопределенности по K_i константе. Аналогичным образом поступаем для поиска границ двумерной области – области по паре констант (рисунок 1).

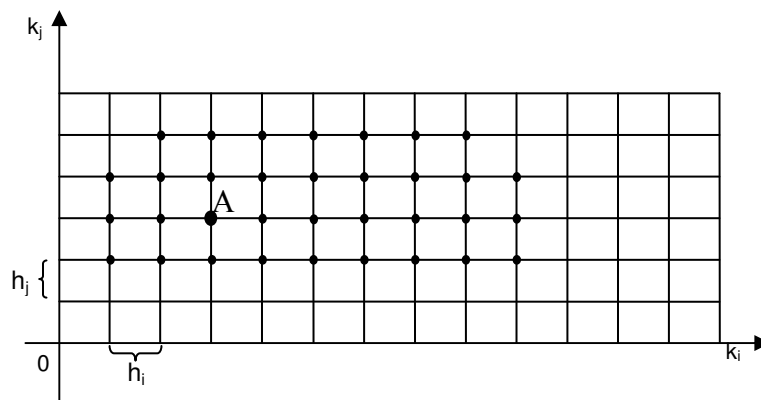


Рис. 1 Схема поиска области неопределенности по двум константам

Оценим объем вычислений. Для определенности сначала будем искать двумерные области. Плоскость по паре констант покрываем сеткой – равномерной прямоугольной. Для нахождения интервалов неопределенности кинетических констант потребуется такой объем вычислений: прямую задачу придется решать $P = C_n^m \cdot q$ раз, где n – общее количество констант, m – количество констант, по набору которых получаем срез области, q – количество узлов сетки, которой покрывается n -мерный куб. Например, количество констант $n=10$, сетка двумерная 1000 на 1000 узлов, тогда потребуется $P = C_{10}^2 \cdot 1000000 = 45000000$ вычислений. Пусть одна прямая

задача (одно вычисление) считается примерно 0,01 секунды. Тогда для поиска двумерной области придется затратить приблизительно 450000 секунд. Это составляет 125 часов или 5,2 дня. А если увеличить количество узлов сетки, если констант будет больше, если увеличить размерность области? Время расчетов увеличится на порядки! Поэтому решено было попробовать перейти на кластеры – использовать многопроцессорные вычисления для существенного ускорения расчетов по времени.

3. Распараллеливание вычислений

Одним из возможных способов проводить крупномасштабные численные исследования является применение параллельных вычислений. Параллельные вычисления можно организовать по-разному, наилучшими для распараллеливания являются задачи, обладающие внутренним параллелизмом, [2] например, включающие параллельные процессы, допускающие распараллеливание по экспериментальной базе.

Наиболее распространенным и легко реализуемым на современных многопроцессорных системах является принцип геометрического параллелизма, который предполагает декомпозицию расчетной области на подобласти по числу процессоров [3].

Распараллеливание вычислительного процесса по экспериментальной базе – решение прямой задачи параллельно для различных наборов начальных данных: при различных температурах, различных начальных концентрациях реагирующих веществ или различных вариантах механизмов протекания реакции. При параллельном доступе к базе данных весь набор узлов кластера разбивается на группы по количеству рассматриваемых наборов экспериментальных данных, и решение для них прямой задачи и задачи поиска интервалов и областей неопределенности осуществляется параллельно, за счет этого достигается наилучшее ускорение вычислительного процесса.

4. Объект исследования

Объектом исследования является разработанная на основании экспериментально обоснованной схемы процесса кинетическая модель реакции циклоалюминирования олефинов триэтилалюминием в присутствии Cp_2ZrCl_2 . [4]

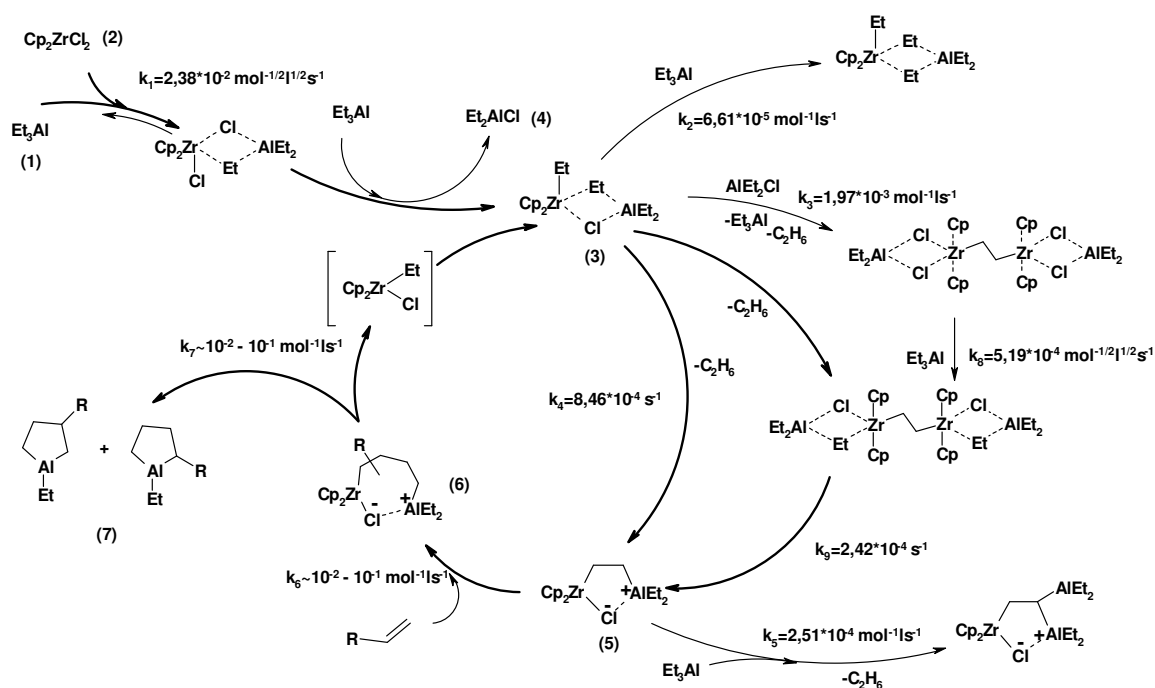
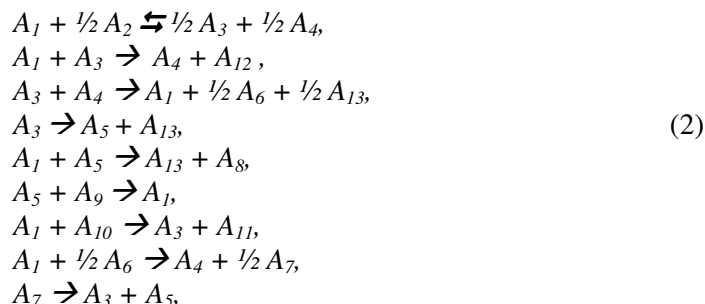


Рис. 2 Схема протекания реакции циклоалюминирования олефинов

Разработанная обобщенная кинетическая модель реакции позволила установить обратимый характер начальной стадии реакции и определить стадии, регулирующие суммарную скорость

всего процесса циклометаллирования, а именно стадии образования пятичленного биметаллического комплекса 5.

Для этой реакции схема химических превращений имеет вид (2), соответствующая ей система уравнений – (3)



где $A_1 = Al(C_2H_5)_3$,

$A_2 = Cp_2ZrCl_2$,

$A_3 = Cp_2Zr(C_2H_5)Cl \cdot Al(C_2H_5)_3$,

$A_4 = ClAl(C_2H_5)_2$,

$A_5 = Cp_2ZrCH_2CH_2Al(Cl)(C_2H_5)_2$,

$A_6 = (Cl)Cp_2ZrCH_2CH_2ZrCp_2(Cl) \cdot 2[ClAl(C_2H_5)_2]$,

$A_7 = (Cl)Cp_2ZrCH_2CH_2ZrCp_2(Cl) \cdot 2[Al(C_2H_5)_3]$,

$A_8 = Cp_2Zr(Cl)CH_2CH[Al(C_2H_5)_2]_2$,

$A_9 = CH_2CHR$,

$A_{10} = Cp_2Zr(Cl)CH_2CHRCH_2CH_2Al(C_2H_5)_2$,

$A_{11} = (C_2H_5)Al(CH_2)_3CHR$,

$A_{12} = Cp_2Zr(C_2H_5)_2 \cdot Al(C_2H_5)$,

$A_{13} = C_2H_6$,

$Cp = C_5H_5$,

R – олефин.

$$\begin{aligned}
 \frac{dX_1}{dt} &= -W_1 - W_2 + W_3 - W_5 - W_7 - W_8; & \frac{dX_8}{dt} &= \frac{1}{2}W_8 - W_9; \\
 \frac{dX_2}{dt} &= -\frac{1}{2}W_1; & \frac{dX_9}{dt} &= W_5; \\
 \frac{dX_3}{dt} &= \frac{1}{2}W_1 - W_2 - W_3 - W_4 + W_7 - W_9; & \frac{dX_{10}}{dt} &= -W_6; \\
 \frac{dX_4}{dt} &= \frac{1}{2}W_1 + W_2 - W_3 + W_8; & \frac{dX_{11}}{dt} &= W_6 - W_7; \\
 \frac{dX_5}{dt} &= \frac{1}{2}W_3 + W_4 + W_5; & \frac{dX_{12}}{dt} &= W_7; \\
 \frac{dX_6}{dt} &= W_4 - W_5 - W_6 + W_9; & \frac{dX_{13}}{dt} &= W_2; \\
 \frac{dX_7}{dt} &= \frac{1}{2}W_3 - \frac{1}{2}W_8;
 \end{aligned} \tag{3}$$

где $W_1 = k_1X_1X_2^{\frac{1}{2}} - k_{10}X_3^{\frac{1}{2}}X_4^{\frac{1}{2}}$, $W_2 = k_2X_2X_3$, $W_3 = k_3X_3X_4$, $W_4 = k_4X_3$, $W_5 = k_5X_1X_5$, $W_6 = k_6X_6X_9$, $W_7 = k_7X_1X_{10}$, $W_8 = k_8X_1X_6^{\frac{1}{2}}$, $W_9 = k_9X_7$ - скорости стадий; X_1, \dots, X_{13} – концентрации веществ A_1, \dots, A_{13} .

Были найдены следующие приведенные значения кинетических констант реакции для соотношения $Cp_2ZrCl_2:AlEt_3=1:10$ (табл. 1)

Таблица 1 Найденные значения кинетических констант

i	1	2	3	4	5	6	7	8	9	10
k _i	165,556	1,121	26,478	4,061	4,843	90,3435	373,5228	26,792	1,986	0,179

В нашем случае измеряемыми компонентами являются $X_2, X_3, X_5, X_7, X_8, X_{12}$, поэтому система неравенств (1) приобретает вид:

$$\begin{aligned} \frac{1}{N} \sum_{j=1}^N |X_{2j}^{\text{э}} - X_{2j}^{\text{п}}| &\leq \varepsilon_2, & \frac{1}{N} \sum_{j=1}^N |X_{7j}^{\text{э}} - X_{7j}^{\text{п}}| &\leq \varepsilon_7, \\ \frac{1}{N} \sum_{j=1}^N |X_{3j}^{\text{э}} - X_{3j}^{\text{п}}| &\leq \varepsilon_3, & \frac{1}{N} \sum_{j=1}^N |X_{8j}^{\text{э}} - X_{8j}^{\text{п}}| &\leq \varepsilon_8, \\ \frac{1}{N} \sum_{j=1}^N |X_{5j}^{\text{э}} - X_{5j}^{\text{п}}| &\leq \varepsilon_5, & \frac{1}{N} \sum_{j=1}^N |X_{12j}^{\text{э}} - X_{12j}^{\text{п}}| &\leq \varepsilon_{12} \end{aligned} \quad (4)$$

где ε_i задаются из условий анализа экспериментальных данных, но не более 15% от экспериментальных значений.

5. Программная реализация

Для решения поставленной задачи использовалась разработанная в Институте нефтехимии и катализа РАН информационно-аналитическая система обратных задач химической кинетики (ИАС ОЗХК) [5].

Вся информация, относящая к эксперименту, сведена в базу данных, которая организована следующим образом: входная информация: количество веществ, стадий, констант, точность численного решения прямой кинетической задачи численным методом Кутта-Мерсона, начальные данные концентраций, набор экспериментальных данных концентраций измеряемых веществ по разным температурам; выходная информация: рассчитанные константы скоростей стадий, энергия активации, левые и правые границы интервалов для пар констант. Для хранения экспериментальных данных кинетических экспериментов в лабораториях ИНК РАН используется формат DAT.

Прямая кинетическая задача решается методом Кутта-Мерсона. Алгоритм решения системы $\dot{x} = f(x)$ данным методом требует пятикратного вычисления правых частей $f(z)$:

$$\begin{aligned} F_1 &= f(x_k) \cdot h; \\ F_2 &= f(x_k + F_1/3) \cdot h; \\ F_3 &= f(x_k + (F_1 + F_2)/6) \cdot h; \\ F_4 &= f(x_k + (F_1 + 3F_3)/8) \cdot h; \\ x_{k+1} &= x_k + (F_1 + 3F_3 + 4F_4)/2; \\ F_5 &= f(x_{k+1}) \cdot h; \\ x_{k+1}^* &= x_k + (F_1 + 4F_4 + F_5)/6, \end{aligned} \quad (5)$$

где $h = \tau_{k+1} - \tau_k$ – шаг интегрирования, x_k – значения переменных в начале интервала интегрирования $[\tau_k, \tau_{k+1}]$. В конце интервала интегрирования τ_{k+1} получается два решения: x_{k+1} – методом Рунге-Кутты с точностью $O(h^4)$ и x_{k+1}^* – методом Кутта-Мерсона с точностью $O(h^5)$.

Систему (3) решали методом перебора вдоль границ при условии выполнения системы (4). Распараллеливание вычислений осуществлялось средствами библиотеки MPI (Message Passing Interface) с использованием языка программирования Fortran. Для проведения расчётов использовался суперкомпьютер Уфимского государственного авиационно-технического университета, имеющий классическую для подобных систем массивно-параллельную архитектуру. Основу суперкомпьютера УГАТУ составляют 266 blade-серверов IBM на базе 532 четырёхъядерных процессоров Intel Xeon.

Параллельная обработка данных основана на следующем принципе: один процессор системы выделяется для сбора данных со всех остальных процессоров и сохранения их в базе данных, также этот процессор (и соответствующий ему процесс), назовем его нулевым, производит считывание всей необходимой входной информации из базы данных и первоначальную рассылку данных остальным процессам для обработки. В результате алгоритм представляет собой следующую последовательность шагов:

- нулевой процесс считывает из базы все необходимые исходные данные и распределяет их по остальным рабочим процессам;
- каждый рабочий процесс рассчитывает область неопределенности для своей пары констант (или одной константы в случае поиска интервала) по описанной выше схеме. В результате получается массив граничных значений области неопределенности;
- каждый рабочий процесс по мере окончания счета производит передачу найденного массива нулевому процессу, который в свою очередь, производит запись полученных данных в базу.

6. Результаты расчетов

С использованием созданного программного комплекса найдены области неопределенности кинетических констант реакции циклоалюминирования олефинов.

Таблица 2 Интервалы неопределенности кинетических констант

№	Константа	Область при $\epsilon=5\%$	Область при $\epsilon=10\%$
1	165,555	[164,451; 169,97]	[154,518; 176,592]
2	1,121	[1,0909; 1,2029]	[1,068; 1,225]
3	26,478	[25,595; 33,715]	[25,595; 36,363]
4	4,061	[3,953; 4,359]	[3,845; 4,576]
5	4,843	[4,617; 5,134]	[4,553; 5,263]
6	90,34	[85,43; 100,06]	[82,3; 101,4]
7	373,523	[370,56; 381,22]	[365,29; 382,44]
8	26,792	[17,326; 29,65]	[10,002; 70,554]
9	1,986	[1,655; 2,105]	[1,43; 2,25]
10	0,179	[0,166; 0,185]	[0,162; 0,185]

Основной вывод по полученным интервалам заключается в том, что начальная точка, вокруг которой находился интервал, не является геометрическим центром этого интервала.

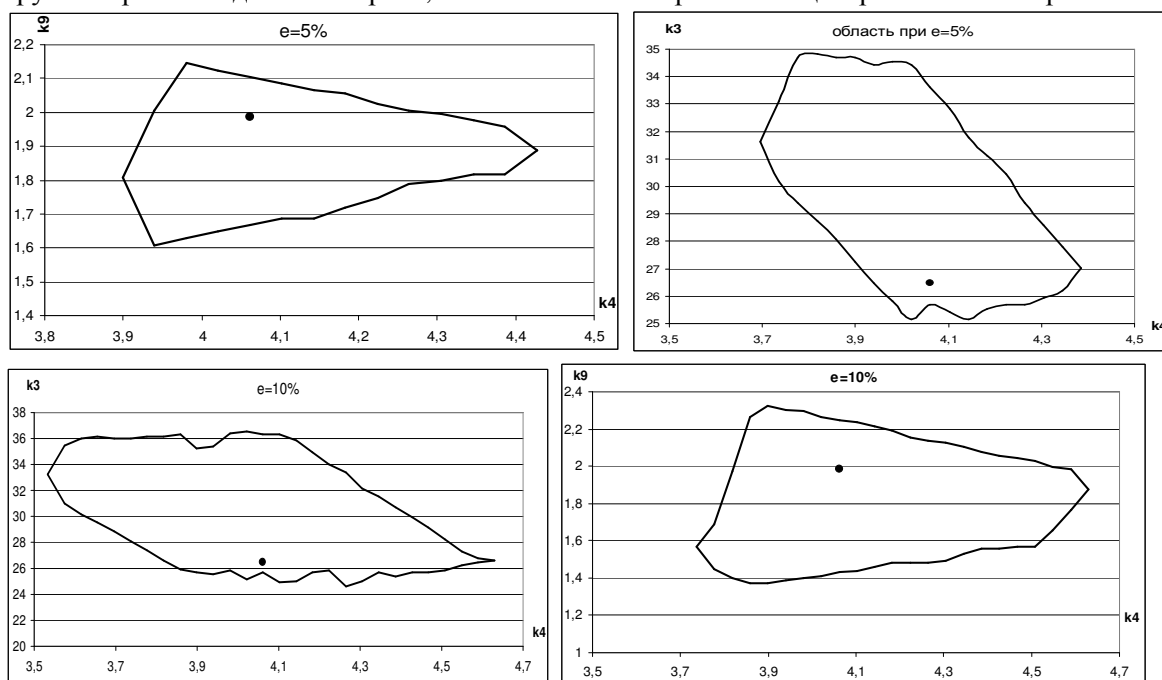


Рис. 3 Фазовые плоскости по некоторым константам при разных ϵ

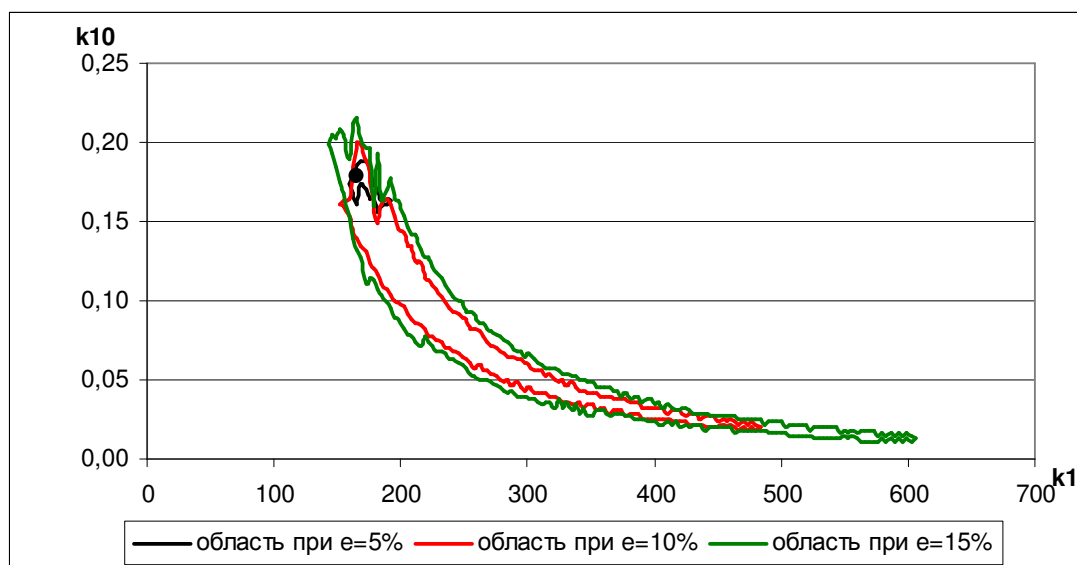


Рис. 4 Сравнение фазовых плоскостей по константам k_1 , k_{10} при разных ε

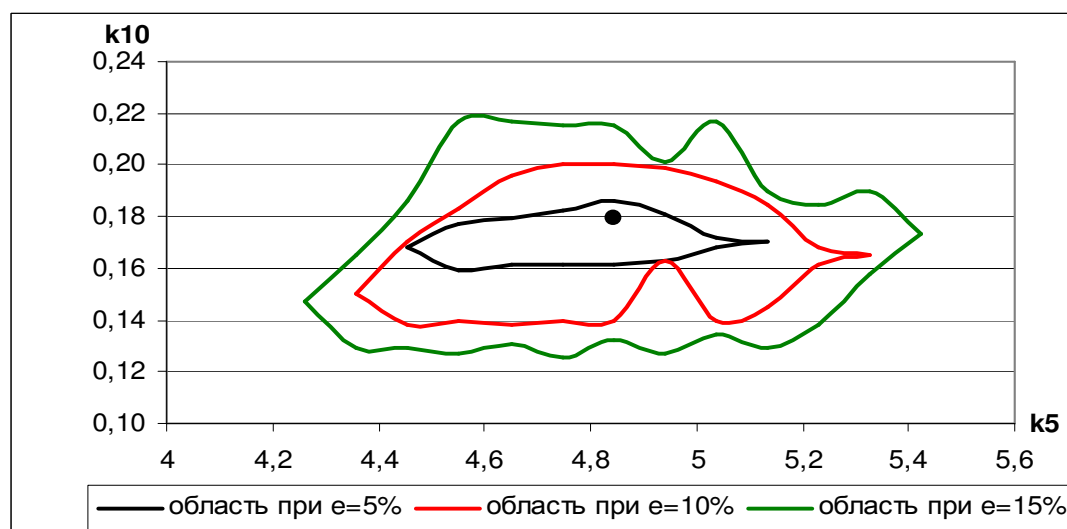


Рис. 5 Сравнение фазовых плоскостей по константам k_5 , k_{10} при разных ε

7. Заключение

Определены кинетические параметры реакции, наилучшим образом описывающие экспериментальные данные, полученные при различных начальных концентрациях компонентов и температурах протекания реакции. Создан метод и алгоритм распараллеливания в обратных задачах химической кинетики. Разработан комплекс прикладных программ для расчета областей неопределенности кинетических параметров с использованием многопроцессорных технологий. Проведен вычислительный эксперимент для анализа данных реакции циклоалюминирования олефинов, построены фазовые плоскости для констант исследуемой реакции.

Литература

1. Канторович Л. В. О некоторых новых подходах к вычислительным методам и обработке наблюдений. // Сиб. мат. журн., 1962, Т. 3. № 5. с. 701-709.
2. Воеводин В.В., Воеводин Вл.В. Параллельные вычисления. СПб: БХВ-Петербург. – 2002. – 608 с.
3. Гергель В.П. Теория и практика параллельных вычислений. – М., 2007. – 423 с.

4. Балаев А. В., Парфенова Л. В., Русаков С. В., Губайдуллин И. М., Спивак С. И., Халилов Л. М., Джемилев У. М. Механизм реакции циклоалюминирования алкенов триэтилалюминием в алюмацклопентаны, катализируемой Cr_2ZrCl_2 .// Доклады Академии Наук, 2001, том 381, №3. с. 364-367.
5. Губайдуллин И.М., Спивак С.И. Информационно-аналитическая система обратных задач химической кинетики // Системы управления и информационные технологии. 2008. №1.1(31). С. 150.