

На правах рукописи



БАВЫКИН ВАСИЛИЙ МИХАЙЛОВИЧ

**ИЗУЧЕНИЕ ЗАВИСИМОСТЕЙ ДЕТОНАЦИОННЫХ
СВОЙСТВ ИНДИВИДУАЛЬНЫХ УГЛЕВОДОРОДОВ
ОТ СТРОЕНИЯ МОЛЕКУЛ**

02.00.03 – «Органическая химия»

02.00.13 – «Нефтехимия»

АВТОРЕФЕРАТ

диссертации на соискание ученой степени

кандидата химических наук

Москва 2008

Диссертационная работа выполнена на кафедре Газохимии Российского государственного университета нефти и газа имени И.М. Губкина и в лаборатории математической химии и компьютерного синтеза Института органической химии имени Н.Д. Зелинского Российской Академии Наук.

Научный руководитель: чл. – корр. РАН, доктор химических наук,
профессор
Липидус Альберт Львович

Научный консультант: кандидат химических наук,
ведущий научный сотрудник
Смоленский Евгений Анатольевич

Официальные оппоненты: доктор химических наук, профессор
Гюльмалнев Агаджан Мирзоевич
доктор химических наук, профессор
Петросян Валерий Самсонович

Ведущая организация: Институт элементоорганических
соединений имени А.Н. Несмеянова
Российской академии наук

Защита состоится «28» октября 2008 г. в 19⁰⁰ часов в аудитории
202 на заседании Диссертационного совета Д 212.200.12 при Российском
государственном университете нефти и газа имени И.М. Губкина по адресу:
РФ, 119991, ГСП-1, В-296, г. Москва, Ленинский проспект, д. 65.

С диссертацией можно ознакомиться в научно-технической библиотеке
РГУ нефти и газа имени И.М. Губкина

Автореферат разослан «26» октября 2008 г.

Ученый секретарь Диссертационного
совета, кандидат технических наук,
доцент



Иванова Л.В.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность проблемы. Задача нахождения зависимостей физико-химических свойств органических соединений от строения их молекул, то есть проблема связи «структура – свойство», является одной из ключевых в органической химии и имеет как теоретическое, так и практическое значение. Во-первых, ее решение необходимо для предсказания свойств неисследованных или еще несинтезированных соединений. Во-вторых, установление связи между структурой молекулы органического соединения и его свойствами дает возможность заменить некоторые трудоемкие эксперименты расчетами, поскольку экспериментальное определение физико-химических свойств органических соединений нередко сопряжено со значительными трудностями, связанными как с самой техникой измерения, так и с получением вещества, его очисткой, нестойкостью, токсичностью, агрессивностью и т.п. Оно требует больших материальных затрат, квалифицированного труда и времени.

В настоящее время проблема связи «структура – свойство» особенно актуальна, т. к. в большинстве случаев требуется не просто получить новое вещество, а синтезировать соединение с определенным набором свойств. Решение этой задачи даже для отдельного, узкого класса соединений, представляется весьма важным. Найденные закономерности можно использовать для систематизации молекул данного класса, прогнозирования свойств гипотетических молекул, поиска новых соединений с заданным набором свойств. Кроме того, установленные закономерности могут оказаться полезными при разработке теории, объясняющей изучаемое явление.

Наличие надежных расчетных методов исследования позволяет предсказывать свойства получаемых вещества (прежде, чем оно синтезировано, а свойство измерено) и тем самым выбрать из многих (ещё не изученных и даже не полученных) соединений те, которые согласно прогнозу, будут удовлетворять поставленным требованиям. Это закладывает научные основы создания новых веществ и материалов с заданными свойствами. Тем не менее, несмотря на то, что по проблеме «структура – свойство» опубликовано множество работ, ее окончательное решение еще не найдено.

Цели и задачи работы. Целями настоящей диссертационной работы стали: установление зависимостей «структура – свойство» для октановых (ОЧ) и цетановых (ЦЧ) чисел алканов и циклоалканов; прогнозирование величин ОЧ и ЦЧ для неисследованных или еще несинтезированных алканов и циклоалканов; описание закономерностей типа «структура – свойство» для ОЧ и ЦЧ алканов и циклоалканов с помощью новых химических дескрипторов (параметр, характеризующий топологическую структуру органического соединения и вычисляемый единственным образом для конкретной молекулы).

В соответствии с поставленными целями основными задачами диссертационной работы явились:

1. Изучение связи ОЧ и ЦЧ алканов и циклоалканов со строением их молекул, вывод расчётных формул и разработка моделей для таких зависимостей;
2. Сравнение полученных результатов расчетов ОЧ и ЦЧ алканов и циклоалканов с экспериментальными данными;
3. Прогнозирование величин ОЧ и ЦЧ для ряда неисследованных или несинтезированных алканов и циклоалканов с помощью разработанных расчетных формул и моделей.
4. Описание зависимостей «структура-свойство» для ОЧ и ЦЧ алканов и циклоалканов с помощью ряда новых топологических индексов и иллюстрация их на конкретных примерах.

Научная новизна. В работе впервые:

1. Выявлены закономерности типа «структура – свойство» для ОЧ алканов, циклопентанов и циклогексанов с помощью разработанных расчётных формул и моделей, позволивших заметно повысить точность расчетов по сравнению с известными ранее методами.
2. Изучены зависимости «структура – свойство» и впервые предложены расчетные формулы для ЦЧ алканов, циклопентанов и циклогексанов.
3. Спрогнозированы значения величин ОЧ (189 углеводов) и ЦЧ (183 углеводорода) для неисследованных или еще несинтезированных алканов и циклоалканов.

4. Описаны зависимости «структура-свойство», с помощью разработанных в настоящей работе уравнений, для вывода которых были предложены новые понятия, связанные с топологической структурой молекул органических соединений: оптимальные топологические индексы (ОТИ), топологические эквиваленты (ТЭ) физико-химических свойств органических соединений и метод обратных функций для решения задач «структура – свойство».

Научно-практическая ценность. Научно-практическая ценность работы определяется тем, что полученные и изученные зависимости ОЧ и ЦЧ от структуры молекул органических соединений, т.е. зависимости «структура – свойство», позволяют прогнозировать величины свойств ОЧ и ЦЧ ещё неисследованных или несинтезированных веществ. Для решения этой задачи были предложены расчетные формулы и новый метод моделирования и прогноза физико-химических свойств, который является универсальным и может применяться не только для расчётов ОЧ и ЦЧ, но и других физико-химических свойств органических соединений.

Методы. Расчёты эмпирических параметров проводились с помощью программного пакета Microsoft Excel (версий 2000, XP, 2003) и ряда оригинальных вспомогательных программ, написанных автором на языке Borland Pascal 7.01.

Апробация работы. Основные результаты диссертационной работы докладывались и были опубликованы в виде тезисов и материалов на: II молодежной конференции ИОХ РАН, Москва, 13-14 апреля 2006 г.; III Международном Симпозиуме «Компьютерное обеспечение химических исследований», Ярославль, 2006 г.; научно-технической конференции «Актуальные проблемы состояния и развития нефтегазового комплекса России», Москва, РГУ нефти и газа им. И.М. Губкина, 29-30 января 2007 г.; X Молодежной конференции по органической химии, ИОХ УНЦ РАН, Уфа, 3-7 декабря 2007 г.; V Международной научно-практической конференции «Новые топлива с присадками», Академия прикладных исследований, Санкт-Петербург, 20-23 мая 2008 г.

Публикации. По теме диссертационной работы опубликованы 5 статей и тезисы 9 докладов на научно-технических, научно-практических конференциях и симпозиумах.

Объем и структура работы. Диссертационная работа состоит из введения (гл. 1), литературного обзора (гл. 2), изложения методов расчёта (гл. 3), основных результатов (гл. 4-5), заключения, выводов, приложений и списка литературы из 158 наименований. Работа изложена на 147 страницах, включает 39 рисунков и 21 таблицу.

Личный вклад автора. Все использованные в диссертационной работе результаты получены автором лично, либо при его определяющем участии как в постановке задач, так и в разработке моделей, проведении численных расчетов и анализе полученных результатов.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

В первой главе (во введении) обоснована актуальность и практическая значимость разработки надежных расчетных методов определения октановых и цетановых чисел углеводородов. Содержатся сведения о задачах, целях, научной новизне, научно-практической ценности, апробации работы и средствах и методах исследования.

Вторая глава содержит обзор литературных данных об исследовании физико-химических свойств индивидуальных углеводородов и моторных топлив, в том числе октановых и цетановых чисел. Приведена информация о лабораторных, стендовых, дорожных, безмоторных и расчетных методах измерения октановых чисел бензинов и индивидуальных углеводородов, а также рассмотрены лабораторные и безмоторные методы оценки воспламеняемости дизельных топлив и индивидуальных углеводородов. Отдельный параграф посвящён состоянию работ в области применения теоретико-графового подхода к изучению проблемы «структура – свойство». Рассматриваются основные понятия теории графов, показано, что структурная формула есть не что иное, как молекулярный граф, где вершины соответствуют атомам, а ребра – химическим связям молекулы. Описана процедура записи

графа в матричном виде, что заметно упрощает работу с последними, вследствие того, что информацию о графе без его графического изображения можно хранить в памяти ЭВМ. Обсуждаются способы решения задач «структура – свойство» с помощью топологических индексов (ТИ), которые являются характеристиками графа, инвариантными относительно изоморфизма, в результате чего они могут быть использованы для кодирования химической информации и для количественного описания химических структур при решении задач «структура – свойство». Рассмотрены преимущества подхода к решению проблемы связи «структура – свойство», основанный на ТИ, а именно: простота и быстрота вычисления ТИ; неэмпирический характер таких индексов; возможность получать их непосредственно из структурной формулы органического соединения. Обоснованы целесообразность и заметные преимущества применения аддитивных схем расчетов и векторного подхода к изучению проблемы «структура – свойство».

В третьей главе подробно описана использованная в настоящей работе методика расчётов. Одной из важнейших задач теоретической химии является моделирование физико-химических свойств углеводородов. В главе описано построение подобных моделей, причем каждой молекуле соответствует некоторый численный параметр, который называют дескриптором. В качестве таких параметров, как правило, используются различные ТИ, которые можно рассматривать как инварианты молекулярных графов, причём атомы водорода не учитываются, т.е. «стёрты». Показано, что любой ТИ и любое физико-химическое свойство можно рассматривать как вектор в N -мерном линейном пространстве (N – число соединений в выборке). В качестве ТИ можно использовать величины

$$g_i = [g_j], \quad (1)$$

где $[g_j]$ – число различных подграфов g_j в графе g_i , соответствующему алкану с номером i ($1 \leq i, j \leq N$). Из чисел $[g_j]$ можно составить матрицу химических структур (МХС):

$$S = \parallel s_{ij} \parallel, \quad (2)$$

столбцы которой и есть вектора $g_j; [g_j]_i = s_{ij}$. Эта матрица имеет треугольный вид, а на главной диагонали стоят одни единицы. Следовательно, $\det(S) = 1$, и векторы g_j являются базисом в этом линейном пространстве размерности N , а любой ТИ, как и любое свойство P , могут быть разложены по этому базису:

$$P = \sum_{j=1}^N a_j g_j, \quad (3)$$

где a_j – коэффициенты разложения.

Величины a_j можно представить, по аналогии с P , в виде вектора A . Тогда матрицу (2) можно записать в матричном виде:

$$P = SA, \quad (4)$$

где вектор P – набор свойства P_N для выборки N , а вектор A – топологический индекс.

Существенным отличием физико-химических свойств от ТИ является то, что в случае ТИ для коэффициентов a_j в (3) справедливо следующее: либо число ненулевых a_j возрастает до конечного предела с ростом N (как для индекса Рандича), либо при этом стремится к бесконечности, но всегда существенно меньше N (индексы Винера, Хосойа). При этом коэффициенты a_j имеют простую и известную зависимость от j или n (для индекса сложности $a_j \equiv 1$, а в случае индекса Винера $a_j \equiv n$ для неразветвлённых подграфов, и $a_j \equiv 0$ в противном случае). Для физико-химических свойств это не так: например, для температуры кипения алканов практически все a_j отличны от нуля. В этом случае можно пренебречь «достаточно малыми» a_j , что позволяет строить модели, способные прогнозировать физико-химическое свойство с заданной наперёд точностью. Индекс

$$I_\nu(\varepsilon) = \sum_{j=1}^N b_j(\varepsilon) g_j, \quad (5)$$

$$b_j = \begin{cases} a_j, & \text{если } |a_j| \geq \varepsilon \\ 0, & \text{если } |a_j| < \varepsilon \end{cases}$$

есть ничто иное, как оптимальный топологический индекс (ОТИ) свойства P с точностью ε .

Доказана теорема о линейности прогноза: если в качестве обучающей выборки используется свойство n -алканов с $n \leq n_0$, а в качестве контрольной – то же свойство n -алканов с $n > n_0$, то вычисленные с помощью ОТИ значения свойства алканов из контрольной выборки линейно зависят от n . Следовательно, ОТИ могут применяться в математическом моделировании физико-химических свойств нормальных и разветвленных алканов только тогда, когда для n -алканов C_nH_{2n+2} , начиная с некоторого $n=n_0$, справедливо соотношение $P(n+1)-P(n)=\text{const}$ (где $P(i)$ – значение физико-химической характеристики для n -алкана C_nH_{2n+2}). В противном случае, если зависимость $P(n)$ монотонна, то ошибка прогноза с ростом n при экстраполяции будет неубывающей.

Показано, что любое физико-химическое свойство органического соединения можно строго упорядочить, предложив математические формулы для количественной оценки «сложности» X и «простоты» Y . Пусть мы имеем экспериментальные данные для свойства P_{exp} , определенные на некотором множестве Q алканов (конечно, определение не зависит от того, рассматриваем мы алканы или же какой-либо другой класс соединений). Назовем «сложностью» свойства P на множестве Q следующую величину:

$$X = \frac{s_0^2}{\Delta_0^2}, \quad (6)$$

где s_0^2 и Δ_0^2 – химическая и обычная дисперсии для свойства P на множестве Q , определенные формулами (7) и (8) соответственно:

$$s_0^2 = \frac{1}{N-1} \sum_{i=1}^N (l_i - 1) \Delta_{0i}^2, \quad (7)$$

$$\Delta_0^2 = \frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{y})^2, \quad (8)$$

где l_i – числа изомеров C_i , y_i – экспериментальные значения, \bar{y} – среднее арифметическое для множества Y . При этом естественно ввести и определение сопряженной величины – «простоты» свойства, а именно:

$$Y = 1 - X = 1 - \frac{s_0^2}{\Delta_0^2} \quad (9)$$

Обобщая вышесказанное, делаем вывод, что «сложность» (β) есть мера группировки экспериментальных значений свойства P_{exp} по изомерам. Таким образом, построенные на известных ТИ модели дают вполне удовлетворительную точность для свойств с величинами $X \leq 0,1$. Модели же для поверхностного натяжения со «сложностью» $X = 0,29$ имеют неудовлетворительную точность, а для температуры плавления, имеющей «сложность» $X = 0,97$, построенные модели фактически лишены смысла.

На рис. 1 - 4 изображены распределения значений ОЧ и ЦЧ алканов и циклоалканов исследуемых выборок по классам изомеров, а также представлены для сравнения распределения величин молекулярной рефракции и молекулярного объема по классам изомеров C_5-C_9 .

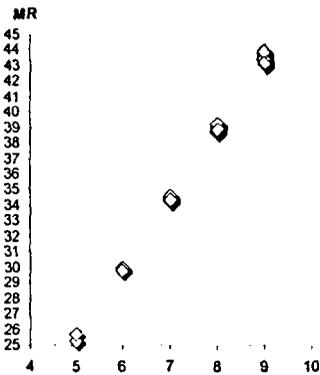


Рис. 1. Распределение величины молекулярной рефракции по классам изомеров C_5-C_9

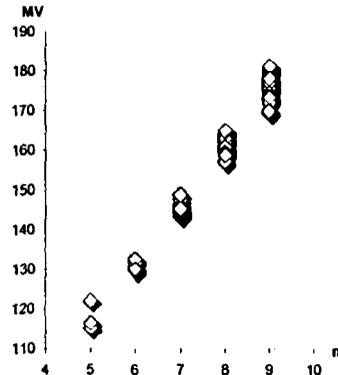


Рис. 2. Распределение величины молекулярного объема по классам изомеров C_5-C_9

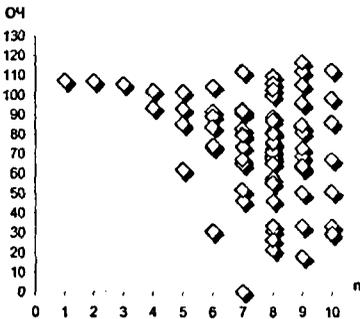


Рис. 3. Распределение величины ОЧ алканов и циклоалканов по классам изомеров C_1-C_{10}

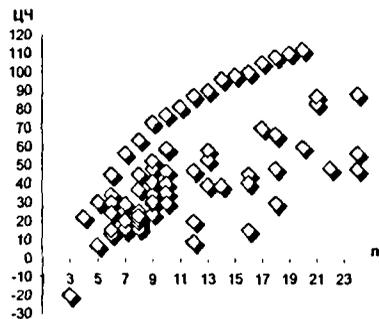


Рис. 4. Распределение величины ЦЧ алканов и циклоалканов по классам изомеров C_3-C_{24}

Эти рисунки позволяют получить наглядное представление о свойствах вновь введенных величин X и Y . Из рис. 1 и 2 видно, что значения молекулярной рефракции и молекулярного объема «плотно» группируются по классам (в данном случае по изомерам), т.е. принадлежность некоторого алкана к определенному классу дает возможность весьма точно предсказать для него по количеству атомов углерода в молекуле величину молекулярной рефракции и молекулярного объема. Поэтому величины сложности этих свойств невелики – 0,0015 и 0,0259 соответственно. Для таких свойств, как ОЧ и ЦЧ, эта картина «размывается» (рис. 3 - 4): значения свойств для разных изомеров начинают «пересекаться». При этом точность определения величины свойства по числу атомов углерода в молекуле уменьшается. Также можно наблюдать, что значения ОЧ и ЦЧ, попавших в один и тот же класс (т.е. имеющих одно и то же значение n), очень сильно различаются между собой, и значение n почти ничего не говорит о возможной величине свойства. Средние значения по изомерам очень мало отличаются от общего среднего.

В результате оказалось, что такие известные свойства, как ОЧ и ЦЧ по «сложности» намного превосходят большинство известных свойств. Проведенный нами расчет «сложности» ОЧ и ЦЧ по формуле (6) дает значения величины сложности $X = 0,78$ и $X = 0,81$ для экспериментальных значений ОЧ алканов и циклоалканов и $X = 0,55$ и $X = 0,49$ – для экспериментальных значений ЦЧ алканов и циклоалканов соответственно. Таким образом, ОЧ и ЦЧ оказались заметно более сложной характеристикой, чем, например, поверхностное натяжение, величина «сложности» для которого составляет $X = 0,29$. Это обстоятельство делает понятным причину того, что расчеты «структура – свойство» для ОЧ и ЦЧ не могли быть проведены с достаточной точностью известными ранее методами.

Для решения задач «структура – свойство» в настоящей работе применен метод топологических эквивалентов. Рассмотрено соотношение «структура – свойство» на примере ОЧ n -алканов (выборка S_1) (здесь и далее – $P_{n, \text{exp}}$), используя $n_1 = n - 1$ (число связей С-С в алкане) в качестве дескриптора.

Эта зависимость хорошо аппроксимируется одной из ветвей гиперболы (рис. 5), которая однозначно определена:

1. координатами точки пересечения своих асимптот – n_0, K_0 ;
2. углом между осью абсцисс и действительной осью гиперболы α (здесь и далее угол отсчитывается от действительной оси гиперболы);
3. длинами действительной $2a$ и мнимой $2b$ осей.

Параметры уравнения второго порядка

$$An_1^2 + Bn_{n,\text{calc}} + CP_{n,\text{calc}}^2 + Dn + EP_{n,\text{calc}} + F = 0, \quad (10)$$

описывающего искомую гиперболу в общем случае, могут быть вычислены, исходя из a, b, n_0, K_0, α по формулам (11-16) с точностью до произвольного ненулевого множителя M :

$$A = M \left(\frac{\cos^2 \alpha}{a^2} - \frac{\sin^2 \alpha}{b^2} \right) \quad (11)$$

$$B = -2M \left(\frac{1}{a^2} + \frac{1}{b^2} \right) \sin \alpha \cos \alpha \quad (12)$$

$$C = M \left(\frac{\sin^2 \alpha}{a^2} - \frac{\cos^2 \alpha}{b^2} \right) \quad (13)$$

$$D = -(BP_0 + 2An_0) \quad (14)$$

$$E = -(Bn_0 + 2CP_0) \quad (15)$$

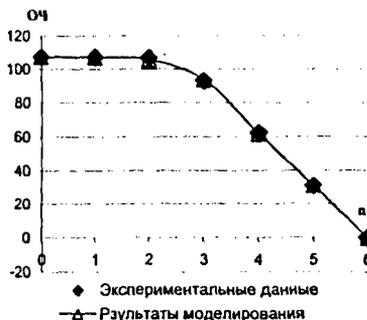
$$F = An_0^2 + Bn_0P_0 + CP_0^2 - M \quad (16)$$

где $P_{n,\text{calc}}$ – вычисленное значение октанового числа алкана C_nH_{2n+2} . Искомая ветвь гиперболы описывается формулами (17-18):

$$P_{n,\text{calc}} = \frac{-Bn_1 - E - \sqrt{(Bn_1 + E)^2 - 4C(An_1^2 + Dn_1 + F)}}{2C} \quad (17)$$

$$n_1 = \frac{-BP_{n,\text{calc}} - D - \sqrt{(BP_{n,\text{calc}} + D)^2 - 4A(CP_{n,\text{calc}}^2 + EP_{n,\text{calc}} + F)}}{2A} \quad (18)$$

Результаты моделирования представлены на рис. 5 и в табл. 1.1.

Рис. 5. Моделирование ОЧ n -алкановТабл. 1.1. Моделирование ОЧ n -алканов

n_1	P_{exp}	P_{calc}	$ P_{exp} - P_{calc} $
0	107,5	107,3	0,2
1	107,1	107,2	0,1
2	105,7	106,7	1,0
3	93,6	93,1	0,5
4	61,8	62,4	0,6
5	31,0	31,1	0,1
6	0,0	-0,3	0,3
R^2	0,99984		
s	0,49603		
$ \Delta_{max} $	1,0		
a	0,9702		
b	1,0016		
n_0	2,5756		
K_0	107,5061		
α	-0,8013		

Преобразуем (18) таким образом, чтобы левая часть уравнения линейно зависела от n_1 и не зависела от $P_{n,calc}$. Получим выражение (19):

$$2An_1 + D = -BP_{n,calc} - \sqrt{(BP_{n,calc} + D)^2 - 4A(CP_{n,calc}^2 + EP_{n,calc} + F)} \quad (19)$$

Введём величину

$$n_{P,exp}(g_i) = -BP_{exp} - \sqrt{(BP_{exp} + D)^2 - 4A(CP_{exp}^2 + EP_{exp} + F)}, \quad (20)$$

полученную из правой части уравнения (19) заменой расчётных значений ОЧ экспериментальными, и назовём её топологическим эквивалентом (ТЭ) ОЧ. Эта величина зависит только от ОЧ, но, в отличие от последнего, успешно аппроксимируется линейной функцией от n_1 , как следует из (19) и (20) (это справедливо и для ТЭ любых других физико-химических свойств). Расчётные значения ТЭ ОЧ $n_{P,calc}(g_i)$, в точности равные левой части уравнения (19), можно преобразовать следующим образом, учитывая, что $n - n_1 = 1$:

$$n_{P,calc}(g_i) = 2An_1 + D = 2An_1 + D(n - n_1) = Dn + (2A - 1)n_1 = I_1 \quad (21)$$

Из уравнения (21) следует, что $n_{P,calc}$ в точности разлагается всего по двум векторам-столбцам МХС, соответствующим графам метана и этана. Таким образом, это значение можно рассматривать как ОТИ физико-химического свойства $n_{P,exp}$ для n -алканов. Так как выразить свободный член через n_1 и n можно для любой линейной функции от n_1 , то это справедливо для ТЭ любых свойств, а не только ОЧ.

Результаты моделирования зависимости $n_{p,exp}(g_i)$ от $n_{p,calc}(g_i)$ для n -алканов представлены на рис. 6 и в табл. 1.2.

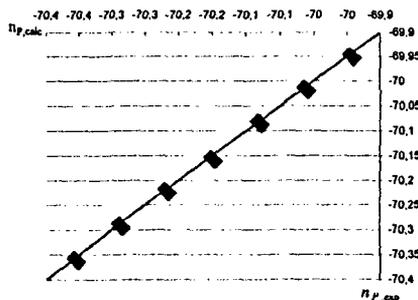


Рис. 6. Моделирование зависимости ТЭ ОЧ $n_{p,exp}(g_i)$ от расчетных значений ТЭ ОЧ $n_{p,calc}(g_i)$ для n -алканов (выборка S_1)

Табл. 1.2. Моделирование ТЭ ОЧ n -алканов (выборка S_1)

n -алкан	$n_{p,exp}$	$n_{p,calc}$	$ n_{p,exp} - n_{p,calc} $
Метан	-70,35693	-70,35686	0,00007
Этан	-70,28866	-70,28826	0,00040
Пропан	-70,21862	-70,21966	0,00104
Бутан	-70,15361	-70,15104	0,00257
Пентан	-70,08021	-70,08246	0,00225
Гексан	-70,01312	-70,01386	0,00074
Гептан	-69,94612	-69,94526	0,00086
R^2			0,99989
S			0,00142
$ \Delta_{max} $			0,00257
a			1,68444
b			1,83242
n_0			3,62115
K_0			107,94588
α			-0,77465

Теперь, исходя из (21), выразим n_i через $n_{p,calc}$

$$n_i = \frac{n_{p,calc}(g_i) - D}{2A} \quad (22)$$

и, подставив это выражение в (17), получим:

$$P_{n,calc} = \frac{-B \frac{n_{p,calc}(g_i) - D}{2A} - E - \sqrt{\left(B \frac{n_{p,calc}(g_i) - D}{2A} + E \right)^2 - C \left(\frac{n_{p,calc}^2(g_i) - D^2}{A} + 4F \right)}}{2C} \quad (23)$$

Далее были проведены расчеты «сложности» ТЭ ОЧ и ЦЧ по приведенным выше формулам, которые дают значения сложности $X = 0,38$ и $X = 0,41$ для ТЭ ОЧ алканов и циклоалканов и $X = 0,21$ и $X = 0,15$ – для ТЭ ЦЧ алканов и циклоалканов соответственно. На рис. 7 и 8 изображено распределение значений ТЭ ОЧ и ЦЧ алканов и циклоалканов исследуемых выборок, что позволяет получить наглядное представление о заметном упрощении задачи.

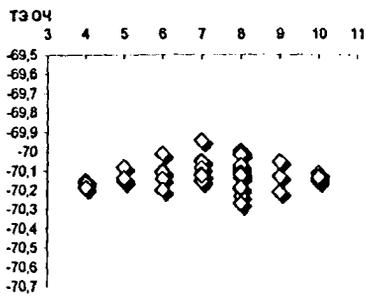


Рис. 7. Распределение величины ТЭ ОЧ по классам изомеров C_4-C_{10}

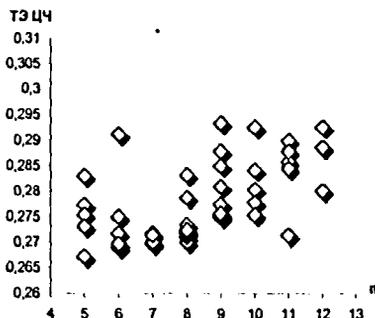


Рис. 8. Распределение величины ТЭ ЦЧ по классам изомеров C_5-C_{12}

Вся выборка из 39 разветвлённых и 32 циклических алканов (в случае с разработкой модели для ОЧ) была разбита на подклассы в соответствии с числом метильных групп в молекулах, если размеры подклассов это позволяли:

- 11 алканов с 3-мя метильными группами (выборка S_2);
- 17 алканов с 4-мя метильными группами (выборка S_3);
- 11 алканов с 5-ю и 6-ю метильными группами (выборка S_4);
- 12 циклопентанов (выборка S_5);
- 20 циклогексанов (выборка S_6).

В случае с ЦЧ, вся выборка из 24 разветвлённого и 31 циклического алкана была разбита на:

- 12 алканов с 3-мя метильными группами (выборка S_2);
- 12 алканов с 4-мя метильными группами (выборка S_3);
- 12 циклопентанов (выборка S_4);
- 19 циклогексанов (выборка S_5).

Разделение циклоалканов на две выборки (как в случае с ОЧ, так и в случае с ЦЧ) обусловлено тем, что граф ни одного циклопентана не содержится в графе какого-либо циклогексана, и наоборот. В случае разделения алканов на подклассы по числу метильных групп сами величины $n_{r,exp}(g_i)$ подвергались разложению только для циклоалканов. В остальных случаях для выборки S_i исследовалась разность между $n_{r,exp}(g_i)$ и $n_{r,calc}(g_i)$, вычисленная по модели для алканов из выборки S_{i-1} .

Четвертая глава содержит результаты моделирования октановых чисел алканов и циклоалканов исследуемой выборки. В решении проблемы линейной зависимости векторов можно применить следующий приём: расчёт $n_{p,exp}(g_i)|_{S_i} - n_{p,calc}(g_i)|_{S_{i-1}}$ проводится для всех найденных ОТИ выборки S_{i-1} , и отбираются те ОТИ, которые дают минимум дисперсии этой величины.

Построенные для выборок S_1 - S_6 модели ТЭ ОЧ на основе ОТИ были объединены в общую модель, при этом каждой из выборок S_5 и S_6 соответствует ТИ L_5 и L_6 – число вхождений в молекулярный граф подграфов циклопентана и циклогексана, соответственно, – которое для соединений, составляющих данную выборку, равно 1, а для соединений, составляющих другие выборки, равно 0. Для выборок S_i ($i = 2, 3, 4$) можно составить индексы L_i , которые равны 0 для соединений из выборок S_j ($j < i$) и не равны 0 для соединений из выборок S_k ($i \leq k \leq 4$) (при $k=i$, $L_i=1$). Индекс L_2 равен вектору, соответствующему изобутану, L_3 – сумме векторов, соответствующих неопентану, 2,3-диметилбутану, 2,4-диметилпентану и т.д., L_4 – сумме векторов, соответствующих 2,2,3-триметилбутану, 2,2,4-триметилпентану, 2,3,4-триметилпентану и т.д. Тогда общая модель для всех выборок может быть представлена формулой

$$n_{p,calc}|_j = \sum_{i=5}^6 L_y I_y + \left(I_{1j} + \sum_{i=2}^4 L_y I_y \right) \left(1 - \sum_{i=5}^6 L_y \right), \quad (24)$$

где $L_y = (L_i)_j$, $I_y = (I_i)_j$.

Теперь, когда величина $n_{p,calc}$ определена для всех 78 исследуемых углеводородов, формула (23) имеет смысл для всех алканов, а не только для нормальных. Становится возможным с её помощью осуществить переход от ТЭ ОЧ к самим октановым числам углеводородов. Необходимо отметить, что разложение (21) индекса I_1 , входящего в итоговое выражение (24), осуществляется с участием параметров A , D уравнения (10), а значит, как следует из выражений (11-16), зависит от всех параметров гиперболы – a , b , n_0 , K_0 , α .

Непосредственная подстановка предшествующих расчётов $n_{p,calc}(g_i)$ для алканов в (24) даёт значение $R^2 = 0,984$ (модель № 1). Это достаточно «плохой» результат, но всё же лучше, чем полученные ранее для той же выборки. Наиболее существенные отклонения от предыдущих значений возникли лишь для параметров, соответствующих молекулярным графам метилэтилпентанов, диэтилпентана и 2,4-диметил-3-этилпентана. Ограничение моделей этими параметрами даёт $R^2 = 0,99669$ (модель № 3). Если дополнить модель коэффициентами в разложении индекса I_2 , то получим $R^2 = 0,99677$ (модель № 4), а если подбирать все коэффициенты, то $R^2 = 0,99715$ (модель № 5). Разница между крайними значениями R^2 моделей составляет всего лишь 0,00046 – достаточно малую величину.

Непосредственная подстановка предшествующих расчётов $n_{p,calc}$ для циклоалканов в (24) даёт значение $R^2 = 0,9958$ (модель № 1с). Вариации по параметрам разложения индексов приводят к незначительному изменению R^2 на $1 \cdot 10^{-9}$, а использование всех параметров даёт $R^2 = 0,9959$ (модель № 2с). Результаты моделирования приведены на рис. 9 и 10.

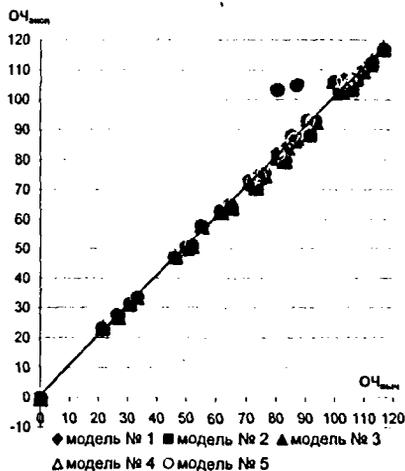


Рис. 9. Моделирование ОЧ алканов

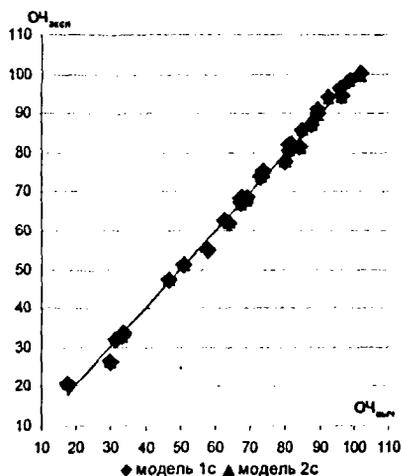


Рис. 10. Моделирование ОЧ циклоалканов

Модель для всех исследованных углеводов ($R^2 = 0,9966$), созданная объединением моделей №№ 3 и 1с, обладает следующими статистическими характеристиками: $R^2 = 0,99660$, $s = 1,53050$, $|\Delta_{\max}| = 6,23228$. Она представлена на рис. 11. Теперь попытаемся найти параметры модели, исходя из минимально

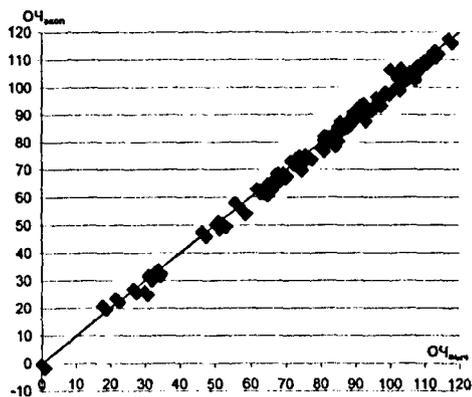


Рис. 11. Общая модель ОЧ предельных углеводов

необходимого количества исходных данных, и осуществить прогноз оставшихся ОЧ. Трудность состоит в том, что для ряда графов, используемых в модели, нельзя подобрать соответствующее соединение, входящее в ту выборку, которую моделирует индексы I_2-I_6 , при построении которых этот граф использован. Это ациклические

подграфы в индексах I_5 , I_6 и графы с числом вершин, меньшим i , в индексах I_i ($2 \leq i \leq 4$). Поэтому приходится использовать ОЧ других соединений (но, конечно, молекулярный граф которых содержит хотя бы один такой подграф). Их количество, разумеется, существенно больше количества таких графов. Поэтому следует осуществить подбор, например, следующим образом. Из контрольной выборки выделялось соединение, для которого отклонение рассчитанного значения от экспериментального максимально по модулю. Затем рассматривались модели, образованные путём замены соединения на любое допустимое из обучающей выборки. Модель, которая обладает наименьшим значением максимального по модулю отклонения рассчитанного значения от экспериментального, исследуется дальше, если эта величина уменьшилась по сравнению с предыдущей моделью, а если нет, то предыдущая модель является окончательной. Обучающая выборка состояла из минимально возможного числа соединений (41 углеводород), контрольная, соответственно, включала 30 соединений.

Такая модель действительно была найдена, причем она обладает следующими статистическими характеристиками: для обучающей выборки $R^2 = 0,99894$, $s = 0,82896$, $|\Delta_{\max}| = 3,0$; для контрольной – $R^2 = 0,99012$, $s = 2,62038$, $|\Delta_{\max}| = 5,6$. Результаты вычислений представлены на рис. 12. С помощью этой модели вычислены значения ОЧ для 189 углеводородов с $n \leq 10$ (алканы, циклопентаны и циклогексаны), которые не синтезированы или ОЧ которых не приведены в литературе.

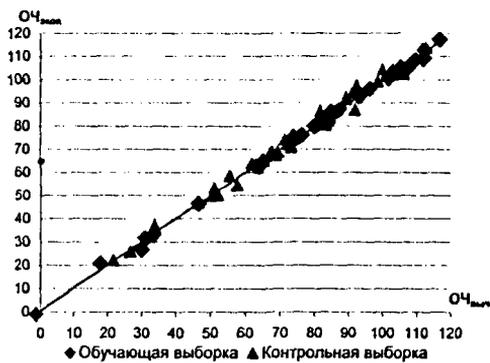


Рис. 12. Экстраполяционная модель ОЧ

С-С в алкане) в качестве дескриптора. На рис. 13 видно, что эта зависимость также хорошо аппроксимируется одной из ветвей гиперболы, которая, как было показано выше, однозначно определена параметрами: a , b , n_0 , K_0 , α .

Как уже отмечалось выше, параметры уравнения второго порядка (10), описывающего искомую гиперболу в общем случае, могут быть вычислены, исходя из a , b , n_0 , K_0 , α по формулам (11-16) с точностью до произвольного ненулевого множителя M . Искомая же ветвь гиперболы описывается формулами (17-18). Результаты моделирования представлены на рис. 13 и приведены в табл. 1.3. Далее были проведены расчеты значений ТЭ ЦЧ для всех выборок по аналогии с расчетами ТЭ ОЧ.

Построенные для выборок S_1 - S_5 модели ТЭ ЦЧ на основе ОТИ были объединены, как и в случае с ОЧ, в общую модель. Для этого поступили следующим образом. Каждой из выборок S_4 , S_5 соответствует уникальный ТИ L_4 , L_5 – число вхождений в молекулярный граф подграфов циклопентана и

Пятая глава содержит результаты моделирования цетановых чисел алканов и циклоалканов, исследуемой выборки. Рассмотрим соотношение «структура – свойство» для ЦЧ 18 n -алканов (выборка S_1) (как и в случае с ОЧ $P_{л,exp}$), используя $n_1 = n - 1$ (число связей

циклогексана, соответственно – который для соединений, составляющих данную выборку, равен 1, а для соединений, составляющих другие выборки, равен 0.

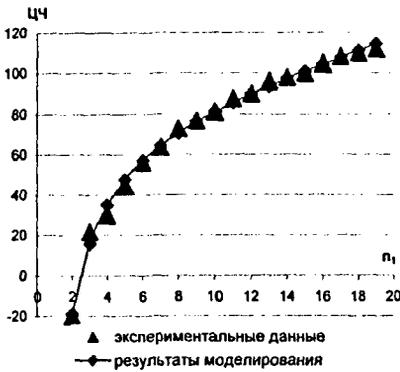


Рис. 13. Моделирование ЦЧ n-алканов

Табл. 1.3. Моделирование ЦЧ n-алканов

n_1	P_{exp}	P_{calc}	$ P_{\text{exp}} - P_{\text{calc}} $	n_1	P_{exp}	P_{calc}	$ P_{\text{exp}} - P_{\text{calc}} $
2	-20,0	-18,8	1,2	11	87,6	85,6	2,0
3	22,0	15,9	6,1	12	90,0	89,8	0,2
4	30,0	35,0	5,0	13	96,1	93,7	2,4
5	44,8	47,7	2,9	14	98,0	97,4	0,6
6	56,3	57,1	0,8	15	100,0	101,0	1,0
7	63,8	64,5	0,7	16	105,0	104,5	0,5
8	73,0	70,8	2,2	17	108,3	107,8	0,5
9	76,9	76,3	0,6	18	110,0	111,1	1,1
10	81,1	81,2	0,1	19	112,0	114,3	2,3
R^2						0,99558	
s						2,31643	
$ \Delta_{\text{max}} $						6,1	
a						8,9451	
b						47,8033	
n_0						-0,1277	
K_0						76,1549	
α						0,1850	

Для выборок S_i ($i = 2, 3$) можно составить индексы L_i , которые равны 0 для соединений из выборок S_j ($j < i$) и не равны 0 для соединений из выборок S_k ($i \leq k \leq 3$) (при $k=i$ $L_i=1$). Индекс \bar{L}_i равен вектору, соответствующему изобутану, \bar{L}_j – сумме векторов, соответствующих неопентану, 2,3-диметилбутану, 2,4-диметилпентану и т.д. Тогда общая модель для всех выборок может быть представлена формулой:

$$\overline{n_{p,\text{calc}}}_j = \sum_{i=4}^5 L_y I_y + \left(I_{1j} + \sum_{i=2}^3 L_y I_y \right) \left(1 - \sum_{i=4}^5 L_y \right), \quad (25)$$

где $L_y = (\bar{L}_i)_j$, $I_y = (\bar{I}_i)_j$.

Теперь, когда величина $n_{p,\text{calc}}(g_i)$ определена для всех 80 исследуемых углеводов, формула (25) имеет смысл для всех алканов, а не только для нормальных. Становится возможным с её помощью осуществить переход от ТЭ ЦЧ к самим ЦЧ углеводов. Непосредственная подстановка предшествующих расчётов $n_{p,\text{calc}}$ для алканов в (10) даёт значение $R^2 = 0,9949$, причём для самых лёгких по молекулярной массе членов выборки – пропана и

бутана – отклонения расчётных значений от экспериментальных максимальны и резко выделяются на общем фоне *n*-алканов (8,0 для бутана, 5,7 для пропана; остальные *n*-алканы – от 2,6 и ниже). Отбрасывание этих значений улучшает R^2 до 0,99677 (модель № 1). Это – достаточно «хороший» результат, тем более что в литературе вообще отсутствуют какие-либо модели ЦЧ индивидуальных углеводородов. Как выяснилось из рассмотрения погрешностей, наибольший вклад в увеличение этой величины внесли 2,3- и 2,4-диметилпентаны. Это означало, что можно обойтись вариациями параметров в разложении погрешностей для индекса I_3 . Это дало $R^2 = 0,99692$ (модель № 2). Из всех вариантов существенное уменьшение значения R^2 наблюдается лишь для параметров, соответствующих молекулярным графам вышеназванных диметилпентанов. Ограничение вариантов этим параметром даёт $R^2 = 0,99682$ (модель № 3). Если дополнить модель коэффициентами в разложении индекса I_2 , то получим $R^2 = 0,99727$ (модель № 4), а если подбирать все коэффициенты, то $R^2 = 0,99736$ (модель № 5). Результаты приведены на рис. 14.

Непосредственная подстановка предшествующих расчётов $n_{p,calc}$ для циклоалканов в (25) даёт значение $R^2 = 0,9836$ (модель № 1с). Использование вариаций по параметрам разложения индексов I_4 , I_5 приводят к значению $R^2 = 0,9840$ (модель № 2с), а использование вариаций по всем параметрам даёт несущественное повышение R^2 на $2 \cdot 10^{-9}$. Результаты представлены на рис. 15.

Модель для всех исследованных углеводородов, созданная объединением моделей №№ 4 и 2с, обладает следующими статистическими характеристиками: $R^2 = 0,9972$, $s = 1,47$, $|\Delta_{max}| = 4,07$. Она представлена на рис. 16.

Также как и в случае с ОЧ была создана экстраполяционная модель, причем она обладает следующими статистическими характеристиками: для обучающей выборки $R^2 = 0,99998$, $s = 0,11680$, $|\Delta_{max}| = 0,5$; для контрольной – $R^2 = 0,99255$, $s = 2,49476$, $|\Delta_{max}| = 6,3$. Результаты вычислений представлены на рис. 17. С её помощью были вычислены значения ЦЧ для 183 углеводородов с $n \leq 10$, которые не синтезированы или ЦЧ которых не измерены.

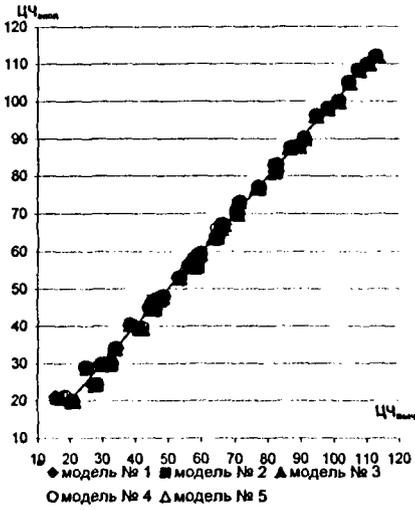


Рис. 14. Моделирование ЦЧ алканов

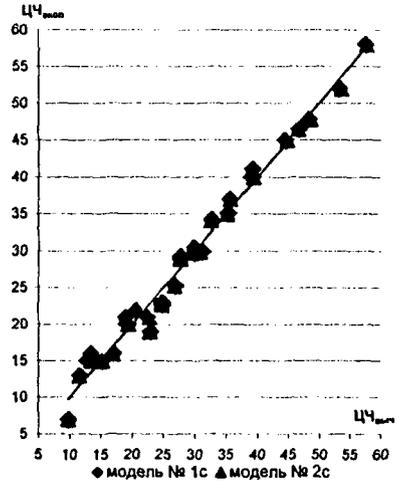


Рис. 15. Моделирование ЦЧ циклоалканов

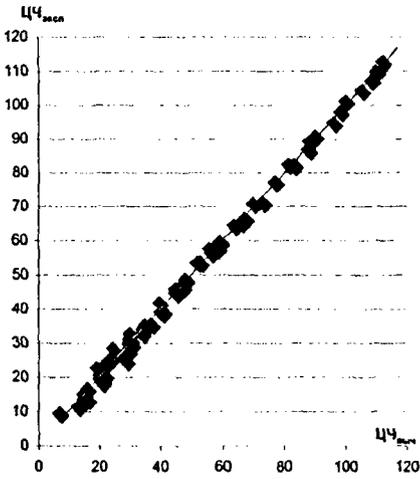


Рис. 16. Общая модель ЦЧ предельных углеводородов

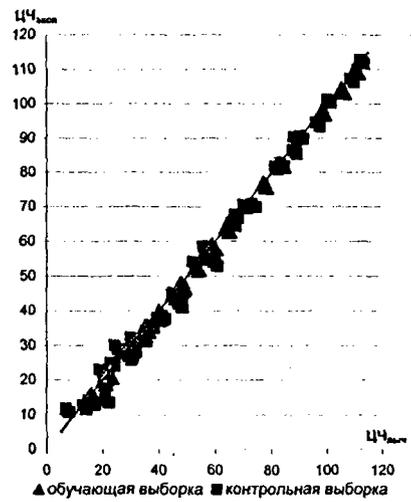


Рис. 17. Экстраполяционная модель ЦЧ предельных углеводородов

Заключение

Таким образом, в представленной работе предложен новый способ расчёта октановых и цетановых чисел углеводородов на основе метода оптимальных топологических индексов (ОТИ), позволивший впервые построить модели, обладающие оптимальными для прогноза статистическими характеристиками. Для использования метода предложено новое понятие – топологический эквивалент свойства (ТЭ), значения которого, в отличие от численного значения самого свойства, всегда линейно зависят от числа атомов углерода (n) для n -алканов, что приводит к несущественности отклонений расчётных значений от истинных, возникающих в результате конечности используемых в обучающих выборках n и в соответствии со сформулированной и доказанной в данной работе теоремой о линейности прогноза.

Высокая эффективность предложенного метода построения расчётных моделей для решения задач «структура – свойство» позволяет предположить возможность его применения для построения моделей других физико-химических свойств не только углеводородов, но и широкого круга различных органических соединений.

Результаты диссертационной работы могут быть использованы для прогноза цетановых и октановых чисел индивидуальных углеводородов с высокой степенью достоверности. Приведенные в приложениях к диссертации результаты прогноза показывают, что некоторые углеводороды могут обладать заметно более эффективными характеристиками в качестве моторных топлив по сравнению с известными и широко применяемыми в настоящее время.

Выводы

1. На основе выявленных зависимостей детонационных свойств от строения молекул разработаны расчетные формулы для ОЧ алканов и циклоалканов, которые значительно превосходят по точности известные ранее схемы.
2. Изучены зависимости и впервые предложены расчетные формулы для ЦЧ алканов, циклопентанов и циклогексанов с высокой прогностической способностью.
3. Рассчитаны и спрогнозированы ОЧ для 189 углеводородов с $n \leq 10$ (61 алкан и 128 циклоалканов) и ЦЧ для 183 углеводородов с $n \leq 10$ (67 алканов и 116 циклоалканов), еще несинтезированных или неисследованных, с помощью разработанных в настоящей работе расчетных формул и моделей.
4. Показана возможность перехода от значений физико-химических свойств органических соединений к их топологическим эквивалентам и обратно.
5. Для решения задач «структура – свойство» и расчета физико-химических свойств органических соединений разработан и апробирован метод обратных функций, в основе которого, в частности, лежит переход от последних к их топологическим эквивалентам.

По теме диссертации опубликованы следующие работы:

1. Смоленский Е.А., Рыжов А.Н., Бавыкин В.М., Мышенкова Т.Н., Лapidус А.Л. Моделирование октановых чисел углеводородов с помощью оптимальных топологических индексов для их топологических эквивалентов // *Изв. АН. Сер. хим.*, 2007, № 9, С. 1619-1632.
2. Смоленский Е.А., Рыжов А.Н., Бавыкин В.М., Чуваева И.В., Лapidус А.Л. Метод обратных функций для решения задач структура - свойство // *ДАН*, 2007, Т. 417, № 3, С. 347-352.
3. Смоленский Е.А., Бавыкин В.М., Рыжов А.Н., Словохотова О.Л., Чуваева И.В., Лapidус А.Л. Расчеты цетановых чисел углеводородов с помощью оптимальных топологических индексов // *Изв. АН. Сер. хим.*, 2008, № 3, С. 451-457.
4. Лapidус А.Л., Бавыкин В.М., Смоленский Е.А., Чуваева И.В. Зависимость цетановых чисел углеводородов от строения молекул // *ДАН*, 2008, Т. 420, № 6, С. 771-776.
5. Лapidус А.Л., Смоленский Е.А., Бавыкин В.М., Мышенкова Т.Н., Кондратьев Л.Т. Модели для расчета и прогнозирования октановых и цетановых чисел индивидуальных углеводородов // *Нефтехимия*, 2008, Т. 48, № 4, С. 277-286.

6. Рыжов А.Н., Бавыкин В.М., Смоленский Е.А., Лapidус А.Л. Моделирование температур кипения алканов с помощью оптимальных топологических индексов (линейное приближение) // *Сборник тезисов докладов 2-ой Молодежной конференции ИОХ РАН, Москва, 2006, С. 168-170.*
7. Рыжов А.Н., Смоленский Е.А., Лapidус А.Л., Бавыкин В.М., Платунов Д.Ю. Нахождение оптимальных топологических индексов для моделирования октановых чисел углеводородов // *Сборник материалов 3-го Международного Симпозиума «Компьютерное обеспечение химических исследований», Ярославль, 2006, С. 22-23.*
8. Lapidus A.L., Smolenskii E.A., Ryzhov A.N., Bavykin V.M., Myshenkova T.N. Modeling of temperature dependences of yields of products of catalytic processes with participation of hydrocarbons // *III International conference «Catalysis: fundamentals and application», Novosibirsk, 2007, Abstracts, V. 2, P. 608-609.*
9. Рыжов А.Н., Бавыкин В.М., Лapidус А.Л., Мышенкова Т.Н., Смоленский Е.А. Теоретические расчеты октановых чисел индивидуальных углеводородов // *Тезисы докладов 7-ой Всероссийской научно-технической конференции «Актуальные проблемы состояния и развития нефтегазового комплекса России», РГУ нефти и газа им. И.М. Губкина, Москва, 2007, С. 238-239.*
10. Бавыкин В.М., Лapidус А.Л., Смоленский Е.А., Мышенкова Т.Н. Моделирование свойств моторных топлив на основе углеводородов // *Сборник тезисов докладов 10-ой Молодежной конференции по органической химии, ИОХ УНЦ РАН, Уфа, 2007, С. 101.*
11. Смоленский Е.А., Бавыкин В.М., Чуваева И.В., Маслова Л.К. Новый подход для нахождения зависимостей «структура-свойство» органических соединений // *Сборник тезисов докладов 10-ой Молодежной конференции по органической химии, ИОХ УНЦ РАН, Уфа, 2007, С. 73.*
12. Лapidус А.Л., Бавыкин В.М., Смоленский Е.А., Чуваева И.В., Мышенкова Т.Н. Моделирование октановых чисел алканов и циклоалканов // *Сборник трудов V Международной научно-практической конференции «Новые топлива с присадками», Академия прикладных исследований, Санкт-Петербург, 2008, С. 63.*
13. Лapidус А.Л., Бавыкин В.М., Смоленский Е.А. Расчетные формулы для вычисления цетановых чисел индивидуальных углеводородов // *Сборник трудов V Международной научно-практической конференции «Новые топлива с присадками», Академия прикладных исследований, Санкт-Петербург, 2008, С. 89.*
14. Смоленский Е.А., Лapidус А.Л., Широков Д.В., Бавыкин В.М. Отклонения от аддитивности октановых чисел смесей углеводородов с высокооктановыми добавками // *Сборник трудов V Международной научно-практической конференции «Новые топлива с присадками», Академия прикладных исследований, Санкт-Петербург, 2008, С. 64.*

Подписано в печать *16.07.08*
Объем

Формат 60x90/16
Тираж *100*

Заказ *504*

119991, Москва, Ленинский просп. ,65
Отдел оперативной полиграфии
РГУ нефти и газа им. И.М. Губкина