

Подсистема «Моделирование клетки»: создание компьютерных моделей функционирования бактериальных клеток, решение задач бактериальной метаболической инженерии и компьютерной поддержки экспериментального дизайна искусственных бактериальных молекулярно - генетических конструкций с заданными свойствами.

Структура документа (оглавление).

1. ЦЕЛЬ И ЗАДАЧИ ПОДСИСТЕМЫ «МОДЕЛИРОВАНИЕ КЛЕТКИ»	3
2. ОПИСАНИЕ КОМПЬЮТЕРНОЙ СИСТЕМЫ «МОДЕЛИРОВАНИЕ КЛЕТКИ» И ДЕТАЛЬНОЕ РУКОВОДСТВО ПО ЕЕ ПРИМЕНЕНИЮ	4
2.1. Программные компоненты компьютерного комплекса «Моделирование клетки»	6
2.1.1. Программная компонента MGSmodeller.....	6
2.1.1.1. Структура и функциональные возможности программной компоненты MGSmodeller.....	6
2.1.1.2. Конструирование/редактирование и расчет математических моделей молекулярно-генетических систем.....	6
Запуск MGSmodeller.....	6
Загрузка и просмотр существующего проекта.....	8
Обновление и сохранение существующего проекта.....	9
Редактирование файлов проекта.....	9
Сборка скрипта.....	10
Расчет прямой задачи для собранного скрипта.....	10
Создание нового проекта.....	13
Работа с проектами вне текущего рабочего пространства (импортирование проекта).....	14
Создание нового скрипта.....	15
Импортирование скрипта в текущий проект.....	16
Создание новых mod и map-файлов в текущем проекте.....	17
Импортирование map-, mod-файлов в текущий проект.....	18
2.1.1.3. Постановка и расчет параметрической обратной задачи (ПОЗ).....	19
Постановка обратной задачи.....	19
Настройка алгоритма расчета обратной задачи.....	22
Расчет обратной задачи.....	24
Постановка и решение задачи оптимального управления.....	26
Постановка задачи управления с помощью мастера.....	26
Запуск на счет задачи управления.....	35
Расчет прямой задачи с управляющими воздействиями.....	37
2.1.2. Программная компонента «METABOL».....	38
2.1.2.1. Структура и функциональные возможности программного компонента «METABOL».....	38
2.2.2.2. Создание компьютерных моделей систем метаболических реакций у бактерий в программной компоненте «METABOL».....	38
Запуск METABOL.....	38
Создание новой модели.....	39
Загрузка существующей модели.....	40
Сборка модели и расчет прямой задачи.....	40
2.1.3. Программная компонента «SETIES».....	42
2.1.3.1. Структура и функциональные возможности программного компонента «SETIES».....	42
2.1.3.1. Создание компьютерных моделей молекулярно-генетических систем в программной компоненте SETIES.....	43
Запуск SETIES.....	44

Регистрация базы данных.....	44
Создание информационной модели генной сети.....	45
..... 46	
Загрузка демо-примеров.....	46
Выборка записи из базы данных.....	46
Сборка информационной модели по make-файлу.....	47
Вызов «Конструктора генных сетей».....	48
Сохранение информационной модели генной сети в журнале экспериментов.....	51
Корректирование входов задачи.....	52
Построение таблицы истинности (ТИ) по заданной в текстовом редакторе дизъюнктивной нормальной форме (ДНФ):.....	52
Построение ДНФ по значениям функции в ТИ.....	53
Расчет динамики в математических моделях.....	53
2.1.4. Программная компонента «STEP+».....	54
2.1.4.1. Структура и функциональные возможности программного компонента «STEP+».....	54
2.1.4.1.1. Создание компьютерных моделей в программном компоненте «STEP+».....	55
Запуск STEP+.....	55
Создание новой модели.....	55
Работа с готовой моделью.....	56
Решение задачи Коши.....	58
Исследование стационарного состояния.....	66
2.1.5. Программная компонента «HGNet».....	70
2.1.5.1. Структура и функциональные возможности программного компонента «HGNet».....	70
2.1.5.1.1. Исследование компьютерных моделей в программном компоненте «HGNet».....	70
Запуск HGNet.....	70
Выбор раздела.....	71
Выбор директории задачи.....	72
Задание исходных данных.....	72
Запуск на расчет.....	73
Опции графики и варьирование параметров.....	74
Анализ результатов.....	75
2.2. Информационные ресурсы для подсистемы "МОДЕЛИРОВАНИЕ КЛЕТКИ".....	76
2.2.1. База данных GeneNet для прокариот.....	76
Организация базы данных GeneNet.....	77
Запросы к базе данных.....	78
Словари.....	78
Источники номенклатурных имен объектов базы данных GeneNet.....	79
Ссылки на внешние ресурсы.....	79
Информационное содержание прокариотической базы данных GeneNet.....	79
2.2.2. База количественных данных и кинетических характеристик (БД Kinet).....	80
Описание работы с базой данных Kinet.....	80
2.2.3. База данных BiotechPro. Решение задач бактериальной метаболической инженерии.....	84
Описание работы с BiotechPro.....	85
Поиск по таблице BiotechProduct.....	86
Примеры типовых запросов для таблицы BiotechProduct.....	89
Поиск по таблице BiotechStrain.....	89
Примеры типовых запросов для таблицы BiotechStrain.....	92
2.2.4. Базы данных GenSensor и ConSensor для дизайна геносенсорных конструкций.....	92
База данных GenSensor.....	93
База данных ConSensor.....	95
3. СПИСОК ЛИТЕРАТУРЫ	97

1. Цель и задачи подсистемы «Моделирование клетки».

Массовое исследование генных сетей (ГС), генетически контролируемых метаболических путей, путей сигнальной трансдукции и других молекулярно-генетических подсистем (МГС) клеток, приводит к исключительно высоким темпам накопления экспериментальной информации, представленной в тысячах публикаций и компьютерных базах данных и описывающих различные аспекты функционирования МГС (Колчанов и др., 2000). Данные исследования в настоящее время вышли на качественно новый уровень в связи с широким распространением техники ДНК-чипов (microarray анализа), позволяющего в одном эксперименте получать информацию о функционировании многих десятков и тысяч генов (Fellenberg et al., 2001, Sherlock et al., 2001). Анализ беспрецедентно огромных объемов экспериментальных данных, отражающих сложные процессы функционирования молекулярно-генетических систем, принципиально невозможен без использования современных информационных технологий и эффективных математических методов анализа данных и моделирования биологических систем и процессов. Для обобщения, систематизации и анализа накопленной информации о генных сетях требуется проведение широкомасштабных теоретических исследований направленных на постижение принципов структурной организации, молекулярных механизмов функционирования, закономерностей эволюции ГС, оценки влияния мутаций на функцию генных сетей, реконструкции ГС на основе экспериментальных данных, для создания искусственных ГС с заданными характеристиками их функционирования (Ananko et al., 2002, 2005). Именно поэтому теоретическое и компьютерное исследование молекулярно-генетических систем приобрело в настоящее время фундаментальное и первоочередное значение. В настоящее время область математического моделирования является динамичным быстро развивающимся направлением исследования закономерностей функционирования живых систем. Арсенал методов, используемых для моделирования молекулярно-генетических систем достаточно широк и включает дискретные, непрерывные, стохастические и комбинированные подходы. Среди них можно выделить Булевы сети (Kauffman, 1993), обобщенные логические сети (Thomas et al., 1995), Байесовские сети (Friedman et al., 2000), динамические Байесовские сети (Ong et al., 2002); сети Петри (Hofstadt, Meineke, 1995), методы моделирования с использованием линейных (Covert et al., 2004) и кусочно-линейных (Edwards, 2000) дифференциальных уравнений; методы моделирования с использованием дифференциальных уравнений, подходы, основанные на применении обыкновенных нелинейных дифференциальных уравнений, в том числе и с запаздывающими аргументами (Smolen et al., 2000; Лихошвай и др., 2003, 2004; Демиденко и др., 2004, Ратушный и др., 2003; Ratushny et al., 2004), стохастическое моделирование (Turner et al., 2004) и др.

ПК «Моделирование клетки» включает в себя современные методы моделирования, расчета и анализа математических моделей молекулярно-генетических и метаболических систем, которые разрабатывались специалистами в области вычислительных методов, теории управления, математического моделирования метаболических путей, математического моделирования МГС.

Основными объектами моделирования в ПК «Моделирование клетки» являются молекулярно-генетические системы клетки (МГС). МГС представляют структурно сложные пространственные объекты, содержащие десятки и сотни (нередко, тысячи) элементов разной природы и сложности: гены и их регуляторные участки; РНК и белки, кодируемые этими генами; низкомолекулярные соединения, различные комплексы между ферментами и их мишенями и т.д. Элементы МГС связываются в единую функциональную сеть посредством сложных нелинейных биохимических процессов синтеза и деградации веществ (Kolchanov et al., 2002). МГС являются открытыми системами, функционирование которых поддерживается

непрерывным поступлением в среду определенных веществ и энергии, а также отводом продуктов деятельности. Функционирование МГС можно характеризовать временными траекториями изменения концентраций некоторой совокупности веществ, принадлежащих ГС.

Важнейшим свойством МГС является способность к изменению состояния (концентрации веществ) в ответ на изменение условий внешней и внутренней среды. Изменение состояния достигается посредством изменения уровня экспрессии определенных групп генов посредством веществ-регуляторов. Регуляторные процессы представляют собой последовательности молекулярных событий (часто достаточно сложных и разветвленных), в которых могут быть одновременно задействованы многие вещества, как поступающие извне (внешние сигналы), так и синтезируемые самой МГС (внутренние сигналы), а также регуляторные участки генов. Ядром МГС являются гены и кодируемые ими РНК и белки, экспрессия которых подвержена взаимному регулированию (Колчанов и др., 2000, Kolchanov et al., 2002). Данные подсети представляют собой регуляторные контуры МГС. Изучение свойств регуляторных контуров является важнейшей задачей, на решение которой ориентирован ПК «Моделирование клетки».

ПК «Моделирование клетки» содержит следующие модули: «конструктор/редактор моделей», «расчет моделей», «обратная задача», «оптимальное управление». Модуль «конструктор/редактор моделей» позволяет создавать и редактировать модели, используя оригинальный стандарт спецификации моделей SiBML. Данный модуль также включает: (i) средства конструирования моделей произвольных МГС с учетом взаимного расположения и ориентации генов в составе геномов, полиаллельности генов, матричного принципа протекания фундаментальных процессов репликации, транскрипции и трансляции, и многокомпарментности исследуемых систем и (ii) средства подготовки моделей для расчета динамики, решения обратных задач и задач оптимального управления. Для расчета динамики моделей в системе реализован метод Гира. Решение обратной задачи осуществляется на основе квазиградиентного и генетического алгоритмов. Задачи управления решаются методом Сеток. Система MGSmodeller реализована в виде java-приложения и снабжена атрибутами специализированного пользовательского интерфейса. Пользовательский интерфейс системы MGSmodeller позволяет иерархически отображать данные, редактировать эти данные и наглядно отображать результаты расчёта и анализа математических моделей МГС.

2. Описание компьютерной системы «Моделирование клетки» и детальное руководство по ее применению

Компьютерная система "Моделирование клетки" предназначена для создания компьютерных моделей функционирования бактериальных клеток, решения задач бактериальной метаболической инженерии и компьютерной поддержки экспериментального дизайна искусственных бактериальных молекулярно-генетических конструкций с заданными свойствами.

Компьютерная система "Моделирование клетки" обеспечивает выполнение следующих функций:

1. конструирование математических моделей молекулярно-генетических систем из моделей элементарных процессов;

2. редактирование моделей путем добавления новых элементов, изменения или удаления фрагментов моделей, изменения условий эксперимента и изменений сценария проведения расчетов;
3. моделирование динамики бактериальных генных сетей с помощью обобщенного химико-кинетического подхода;
4. моделирование стационарного распределения метаболических потоков в бактериальной клетке;
5. использование обобщенных функций Хилла для моделирования нелинейных эффектов функционирования промоторов и аллостерических механизмов регуляции функции ферментов.
6. численное исследование динамики переходных процессов математических моделей молекулярно-генетических систем и устойчивости предельных состояний математических моделей молекулярно-генетических систем.
7. поиск оптимальных параметров математических моделей молекулярно-генетических систем, обеспечивающих наилучшее соответствие динамики компонентов моделей экспериментальным данным;
8. решение задач управления динамикой функционирования молекулярно-генетических систем, в частности:
 - поиск управляющих воздействий на динамику генной сети с целью перевода ее из одного стационарного состояния в другое;
 - оценка побочных эффектов и зоны риска при управляющих воздействиях на динамику генной сети.
9. конструирование математических искусственных генных сетей на основе математических моделей бактериальных генетических элементов;
10. исследование динамики функционирования искусственных генных сетей с заданными наборами генетических элементов и графами регуляторных связей;
11. информационную поддержку экспериментального дизайна гено - сенсорных устройств на основе бактериальных систем.

Компьютерная система «Моделирование клетки» состоит из пяти программных (MGSmodeller, METABOL, SETIES, STEP+, HGNET) и пяти информационных компонентов: базы данных GeneNet, KiNET, BiotechPro, GenSensor, ConSensor. Вызов доступных ресурсов осуществляется с общей панели. В дальнейшем работа с вызванным ресурсом осуществляется независимо от других ресурсов системы. В связи с этим в данном обучающем курсе руководство по применению каждого компонента (программного и информационного) дается в отдельном разделе.

Понимание текста размещенного в разделе 2 подразумевает, что пользователь ознакомился с методами моделирования и со стандартами спецификации данных, применяемых в программных и информационных компонентах компьютерной системы «Моделирование клетки», подробное описание которых дано в руководстве пользователя (АСНИ-01).

2.1. Программные компоненты компьютерного комплекса «Моделирование клетки»

2.1.1. Программная компонента MGSmodeller

2.1.1.1. Структура и функциональные возможности программной компоненты MGSmodeller.

Программная компонента MGSmodeller предназначена для создания и анализа компьютерных моделей функционирования подсистем бактериальных клеток.

MGSmodeller позволяет создавать и редактировать модели, написанные с использованием обобщенного химико-кинетического подхода и внутреннего стандарта спецификации моделей SiBML, в том числе предоставляет: (i) средства конструирования моделей произвольных МГС с учетом взаимного расположения и ориентации генов в составе геномов, матричного принципа протекания фундаментальных процессов репликации, транскрипции и трансляции, и многокомпарментности исследуемых систем и (ii) средства подготовки моделей для расчета динамики, решения обратных задач и задач оптимального управления.

MGSmodeller обеспечивает решение задачи Коши на заданном интервале времени. Для расчета динамики моделей в MGSmodeller реализован метод Гира.

MGSmodeller снабжен средствами поиска оптимальных значений параметров моделей. Решение обратной задачи осуществляется на основе квазиградиентного и генетического алгоритмов.

MGSmodeller позволяет формулировать и решать задачи оптимального управления на основе метода Сеток.

Понимание текста размещенного в разделе 2.1.1. подразумевает, что пользователь ознакомился с обобщенным химико-кинетическим методом моделирования (ОХКММ) и со стандартом спецификации моделей SiBML, применяемых в программном комплексе "Моделирование клетки". Подробное описание ОХКММ и стандарта SiBML дано формат файлов см. "ОИОС01". Подробное описание руководства пользователя дано в "АСНИ-01".

2.1.1.2. Конструирование/редактирование и расчет математических моделей молекулярно-генетических систем

Запуск MGSmodeller

Запустите исполняемый файл modeller.exe в директории, в которой размещен MGSmodeller.

Появляется диалоговое окно выбора рабочего пространства (Рис. 1).

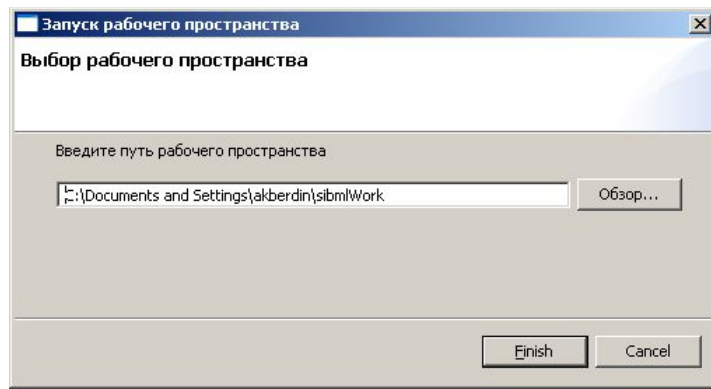


Рисунок 1. Диалоговое окно выбора рабочего пространства

Ввод пути к рабочему пространству, где будут находиться все созданные проекты и настройки пути, можно осуществлять вручную, либо используя вспомогательный диалог, который активизируется после нажатия кнопки “Обзор” (Рис. 2).



Рисунок 2. Диалоговое окно выбора директории рабочего пространства

Для загрузки проектов, содержащих обучающие примеры, выберите поддиректорию samples в директории, в которой размещен MGSmodeller (Рис. 3).

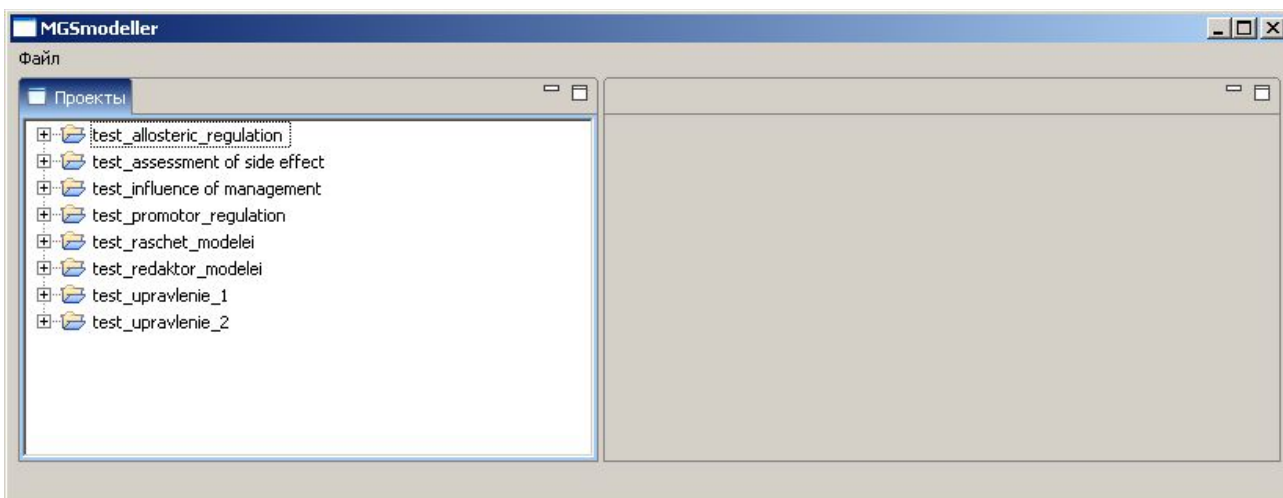


Рисунок 3. Главное рабочее окно MGSmodeller с активной закладкой проектов, содержащих обучающие примеры.

Загрузка и просмотр существующего проекта

Выберите проект «test_redaktor_modelei», откройте поддиректории проекта и кликните на интересующий Вас файл дважды. Содержимое файла откроется в окне редактора. Вы можете открыть несколько файлов одновременно (Рис. 5).

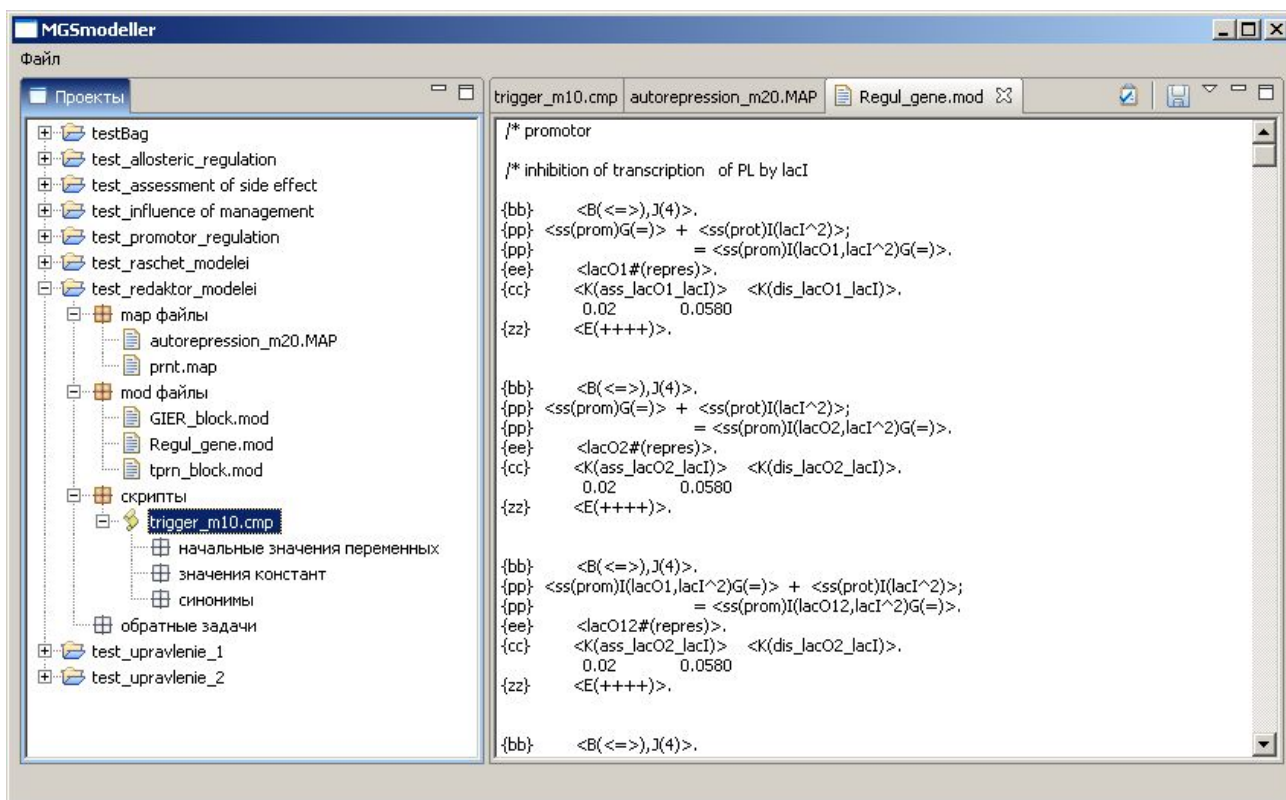


Рисунок 4. Пример окна просмотра файлов в проекте, с визуализацией mod файла.

Обновление и сохранение существующего проекта

Операция обновления проекта необходима в следующих случаях:

- выполнены операции по постановке и расчету прямой, обратной и задачи управления
- в папку где физически хранится проект скопированы tar, mod или str файлы
- из папки где физически хранится проект удалены tar, mod или str файлы

Операция сохранение выполняется для того, чтобы все изменения внесенные Вами были доступны при следующем запуске системы.

Для того чтобы обновить, кликните на имени интересующего Вас проекта правой кнопкой мыши. В выпадающем списке доступных действий выберите «Обновить проект». Для того чтобы сохранить, кликните по имени интересующего Вас проекта правой кнопкой мыши. В выпадающем списке доступных действий выберите «Записать проект»(Рис. 5).

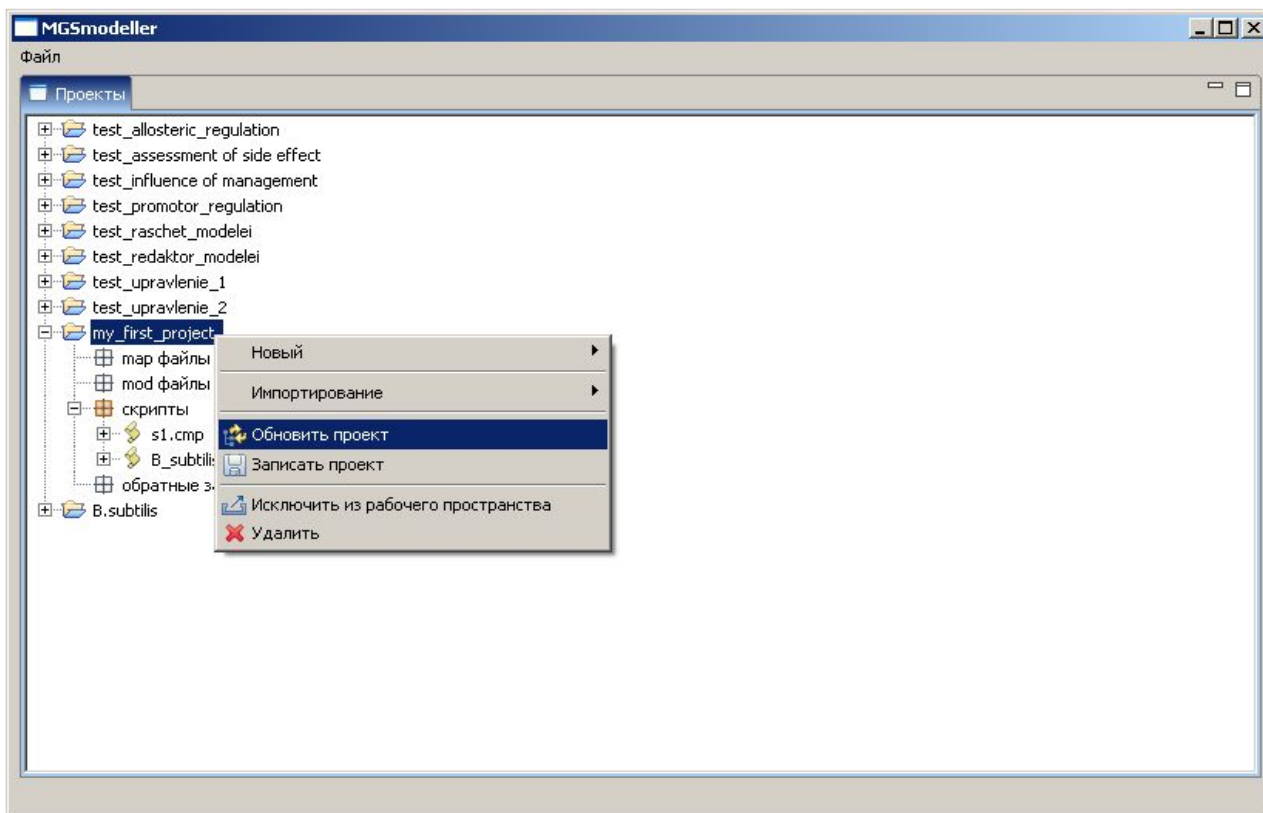




Рисунок 5. Пример окна действий обновления и сохранения проекта.

Редактирование файлов проекта

Откройте двойным кликом левой клавишей мыши содержимое файла проекта в окне редактирования. После этого Вы можете приступить к редактированию файла с учетом принятых для каждого файла стандартов спецификаций (форматы файлов см. "ОИОС01"). Для проверки синтаксиса файла необходимо нажать кнопку  в правом верхнем углу окна редактора. Для сохранения файла необходимо нажать кнопку  в правом верхнем углу окна редактора.

Сборка скрипта

Для сборки скрипта закройте все открытые файлы. Нажмите правой кнопкой мыши на имя скрипта в дереве проекта (скрипту соответствует файл с расширением `cmp`) и выбрать в выпадающем меню пункт «Собрать скрипт» (Рис. 6).

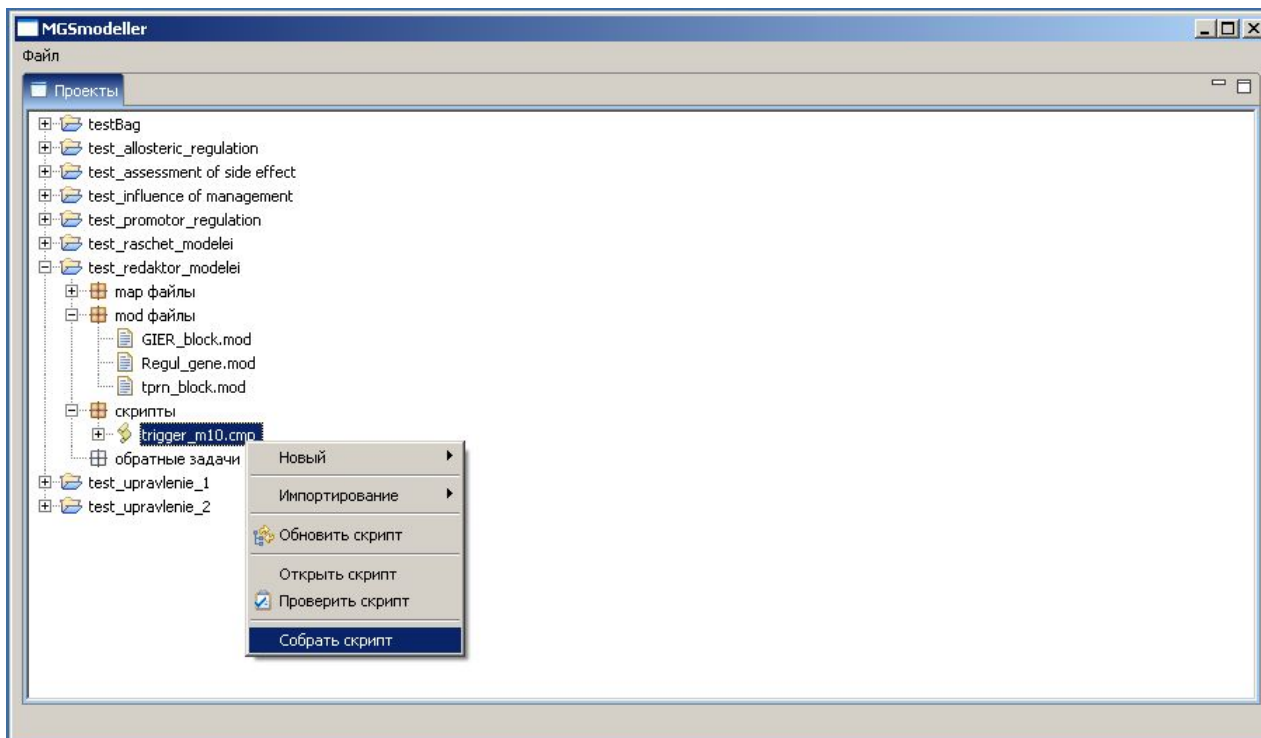
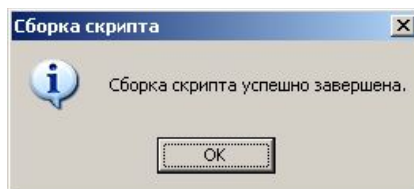


Рисунок 6. Сборка скрипта.

Сборка скрипта может занять некоторое время. Дождитесь появления следующего окна:



Нажмите ОК. Сборка скрипта успешно завершена.

Расчет прямой задачи для собранного скрипта

После сборки скрипта, его поддерево дополняется необходимыми для расчета файлами. В файле `*_all.vpr` (где символом `*` обозначено названия скрипта) представлены все переменные скрипта. В файле `*_all.vrc` (где символом `*` обозначено названия скрипта) представлены все константы скрипта (Рис. 7). Скрипт формально готов к расчету прямой задачи.

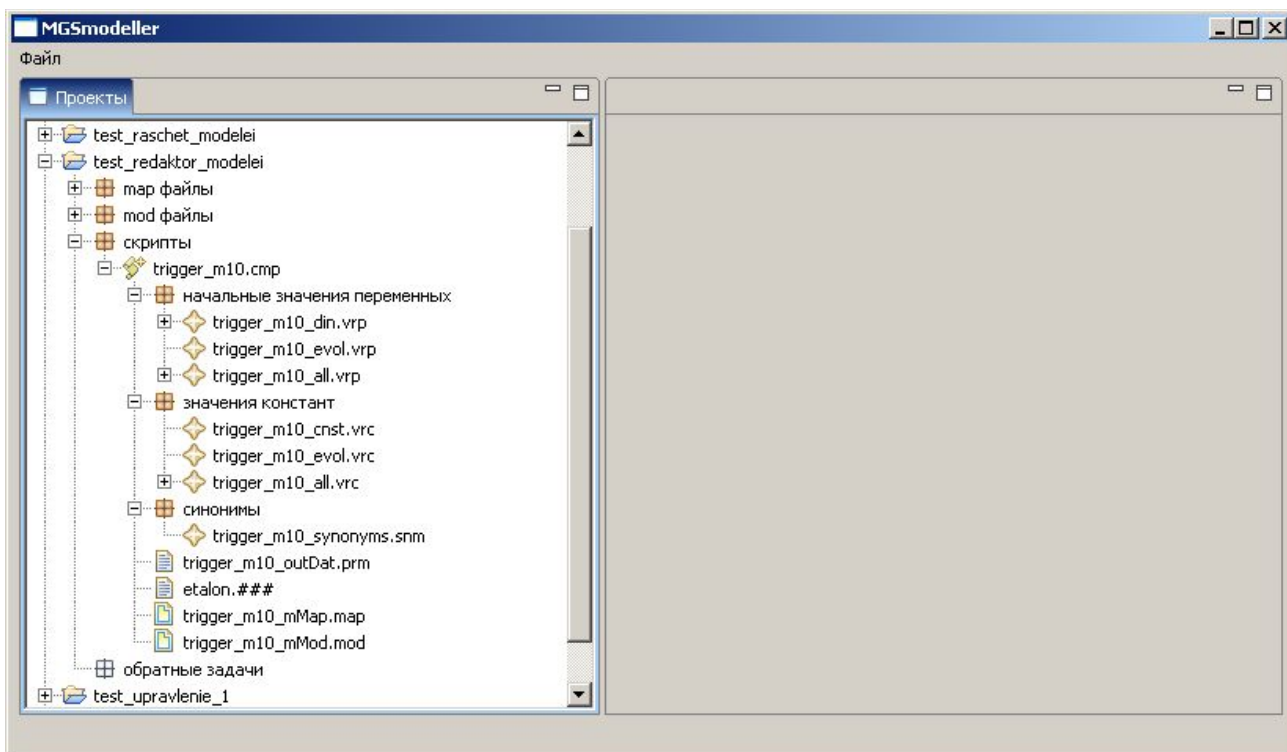




Рисунок 7. Скриншот окна после сборки скрипта.

При необходимости пользователь может на данном этапе изменить начальных значения переменных в файле *_din.vrp (где символом "*" обозначено названия скрипта), а также изменить значения констант в файле *_cnst.vrc (где символом "*" обозначено названия скрипта). Для редактирования соответствующего файла откройте его двойным кликом левой клавишей мыши в окне редактора. Редактирование файлов осуществляется с учетом принятых для каждого файла стандартов спецификаций (форматы файлов см. "ОИОС01"). Для проверки синтаксиса файла необходимо нажать кнопку  в правом верхнем углу окна редактора. Для сохранения файла необходимо нажать кнопку  в правом верхнем углу окна редактора.

Для запуска расчета прямой задачи скрипта закройте все открытые файлы. Нажмите правой кнопкой мыши на имя скрипта в дереве проекта и выбрать в выпадающем меню пункт «Прямая задача»/«Запустить на счет» (Рис. 8).

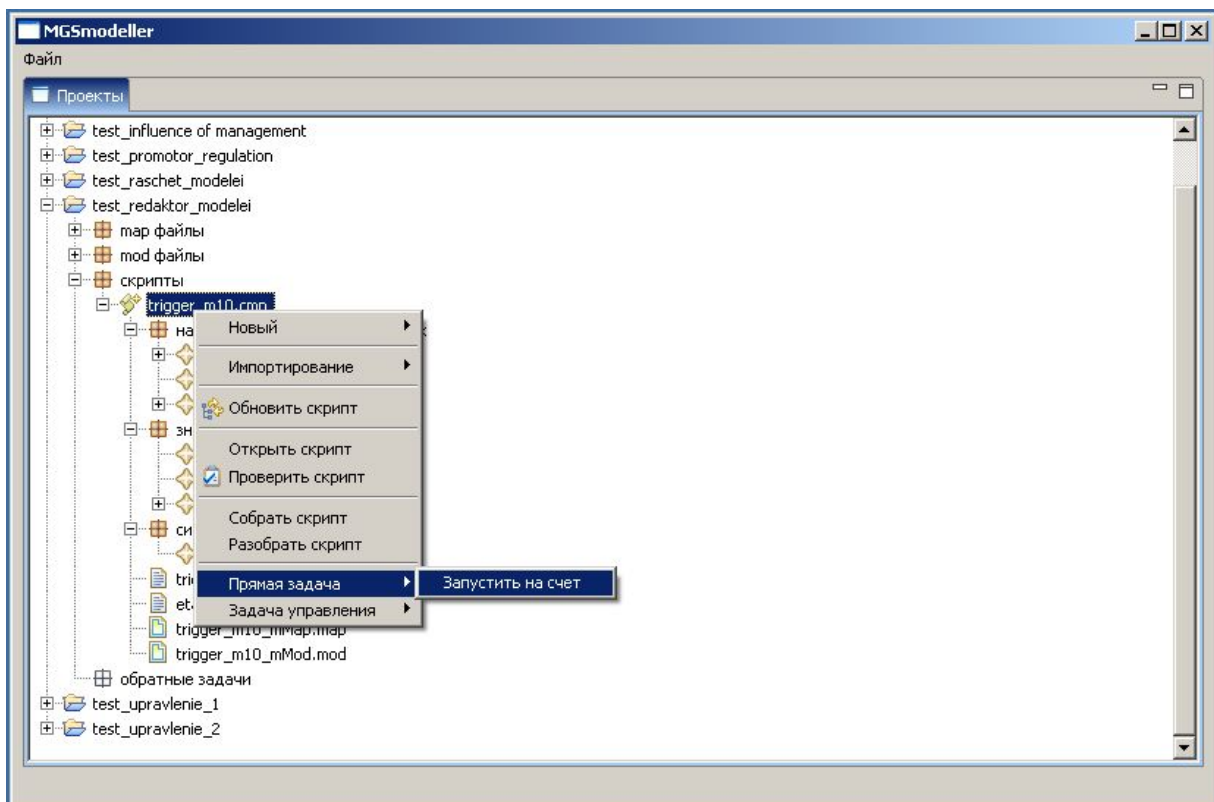
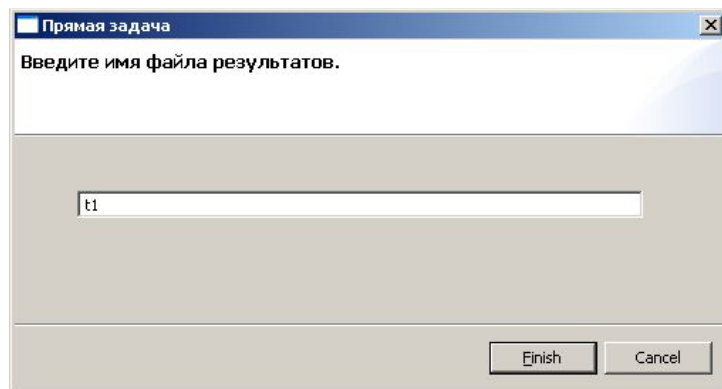
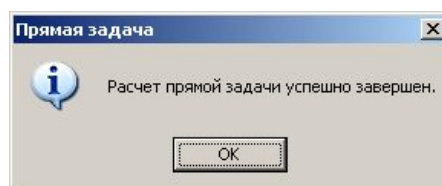


Рисунок 8. Запуск расчета прямой задачи скрипта.

В появившемся диалоге введите имя, под которым будут сохранены в скрипте результаты решения прямой задачи. После нажатия кнопки “Finish” запускается расчет прямой задачи.



Расчет прямой задачи скрипта может занять некоторое время. Дождитесь появления следующего окна:



Нажмите ОК. Расчет прямой задачи скрипта успешно завершен.

Для просмотра результатов расчетов скрипта нажмите правой клавишей мыши на имя результатов расчета прямой задачи в поддереве скрипта «результаты расчетов» и в выпадающем меню пункта «Открыть» выберите любой из предложенных пунктов (Рис. 9).

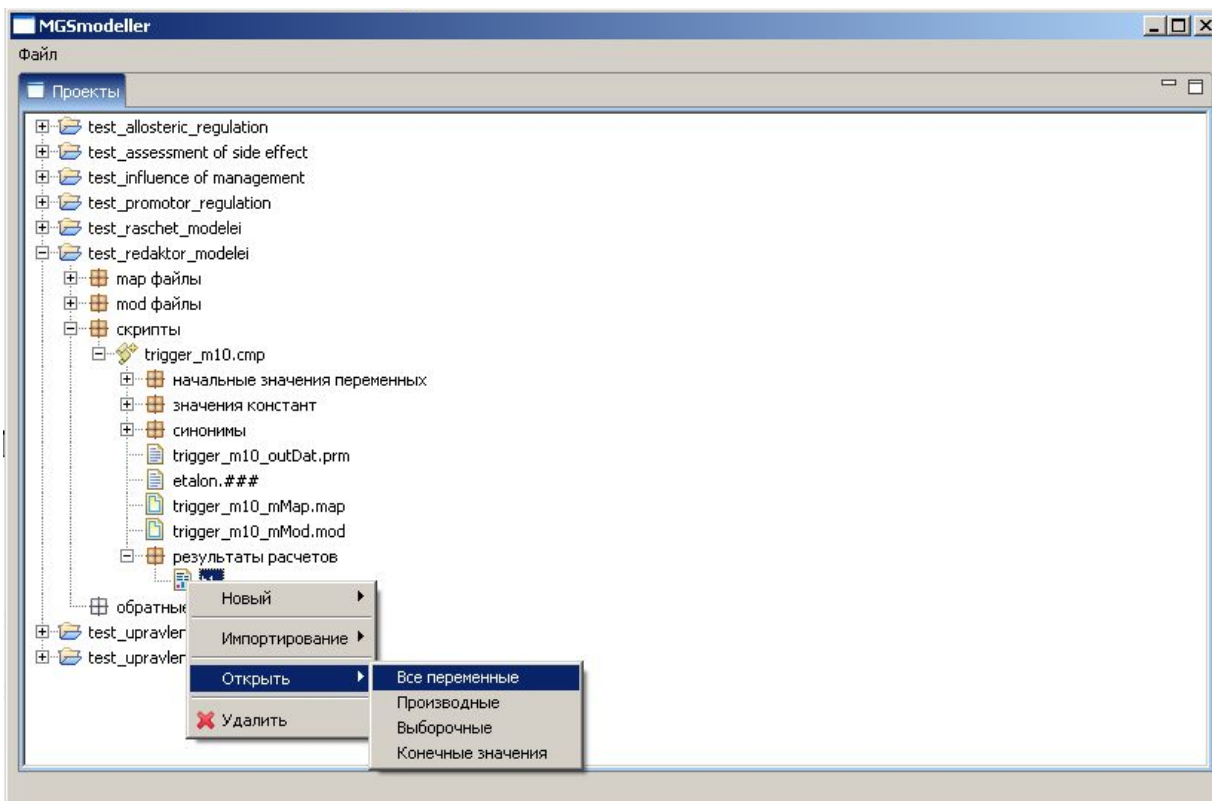


Рисунок 9. Просмотр результатов расчета прямой задачи скрипта.

Создание нового проекта

Для создания нового проекта нажмите правой кнопкой мыши в области вкладки «Проекты» и выбрать в выпадающем меню пункт «Новый»/«Проект...» (Рис. 10).

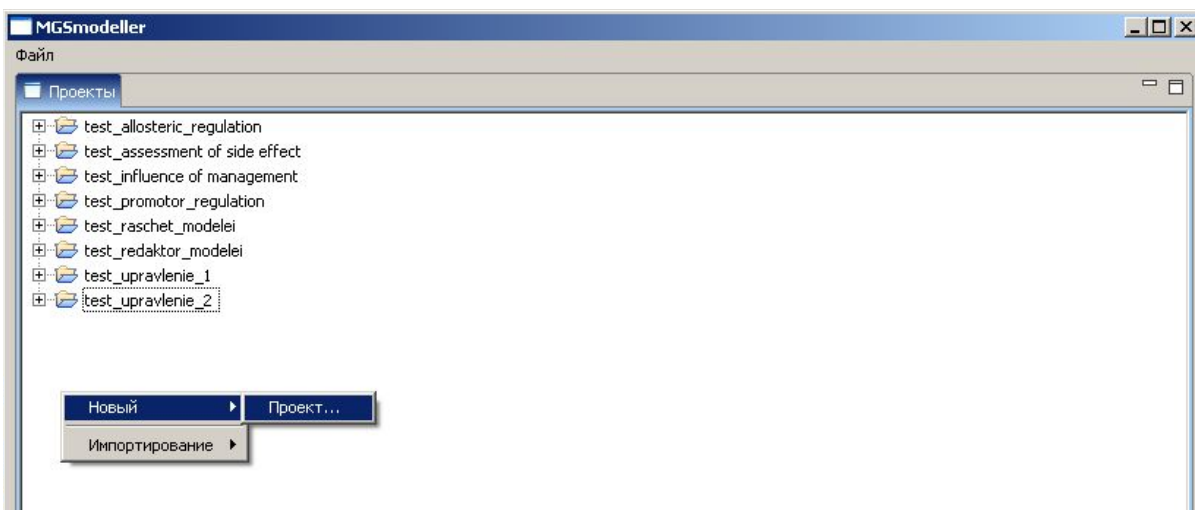
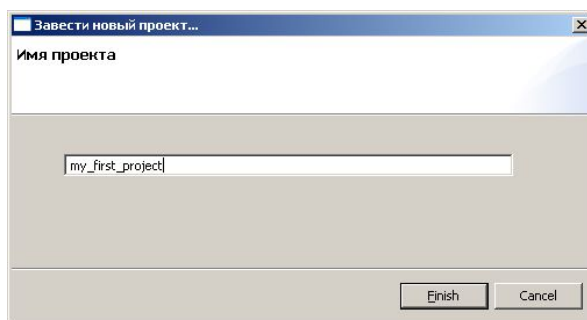


Рисунок 10. Создание нового проекта.

В появившемся диалоге введите имя нового проекта.



После нажатия кнопки “Finish” создается новый пустой проект.

Работа с проектами вне текущего рабочего пространства (импортирование проекта)

Операция импортирования проекта в MGSmodeller является операцией открытия проекта лежащего вне текущего рабочего пространства. Для импортирования проекта нажмите правой кнопкой мыши в области вкладки «Проекты» и выбрать в выпадающем меню пункт «Импортирование»/«Проект...» (Рис. 11).

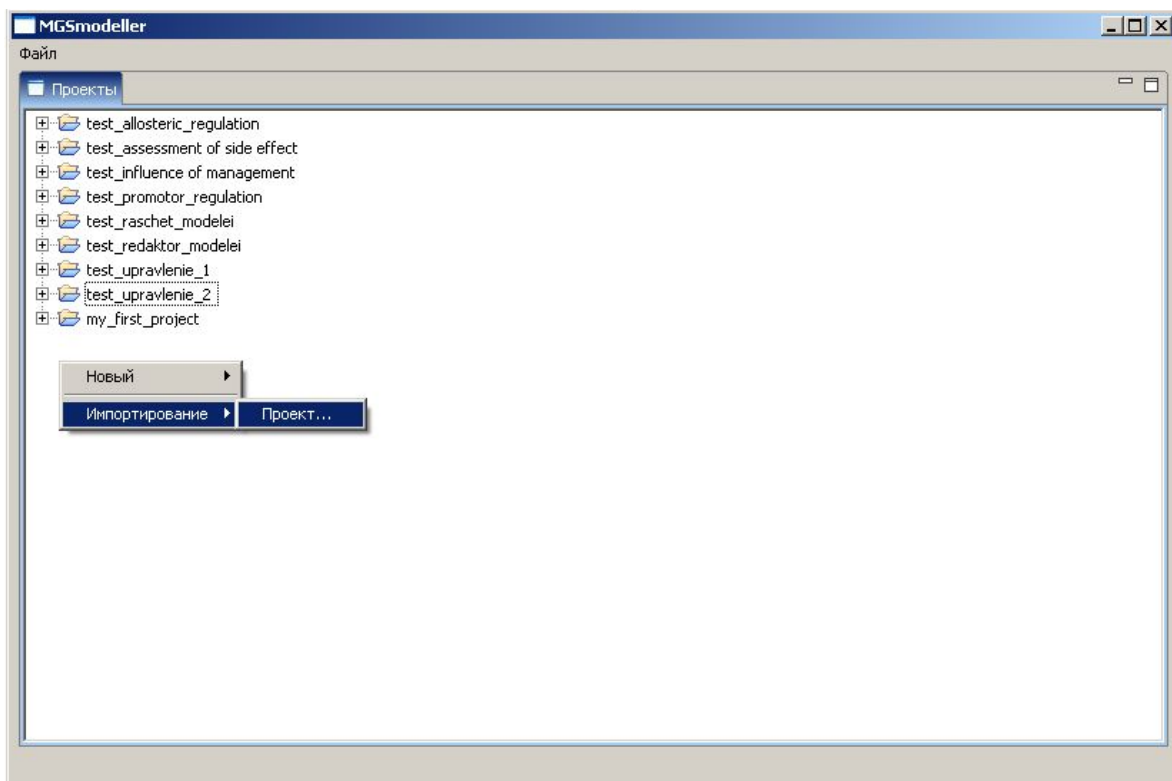
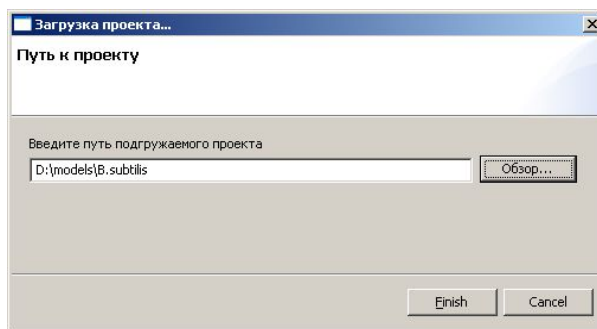


Рисунок 11. Импортирование проекта.

В появившемся диалоге введите путь к импортируемому проекту.



После нажатия кнопки "Finish" проект импортируется в рабочее пространство.

Создание нового скрипта

Для создания нового скрипта нажмите правой кнопкой мыши на поддерево «скрипты» в дереве текущего проекта и выберите в выпадающем меню пункт «Новый»/«скрипт» (Рис.12).

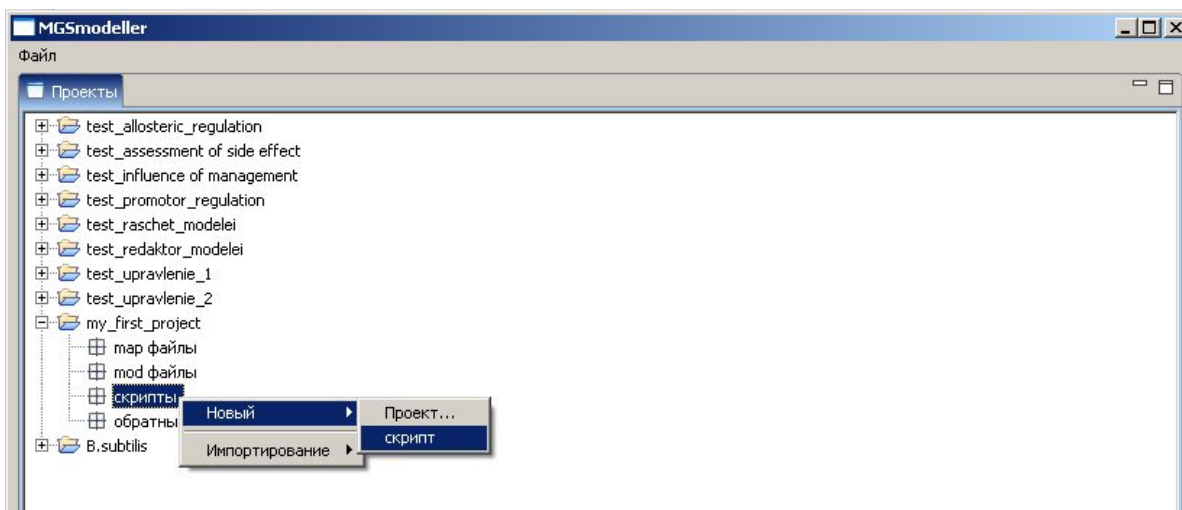
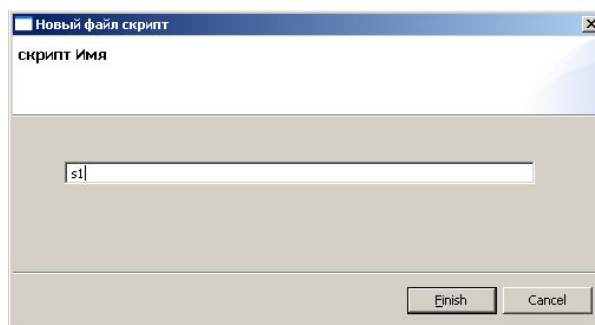


Рисунок 12. Создание нового скрипта в текущем проекте.

В появившемся диалоге введите имя нового скрипта.



После нажатия кнопки "Finish" создается новый пустой скрипт.

Импортирование скрипта в текущий проект

Для импортирования скрипта в текущий проект нажмите правой кнопкой мыши на подерево «скрипты» в дереве текущего проекта и выберите в выпадающем меню пункт «Импортирование»/«скрипт» (Рис.13).

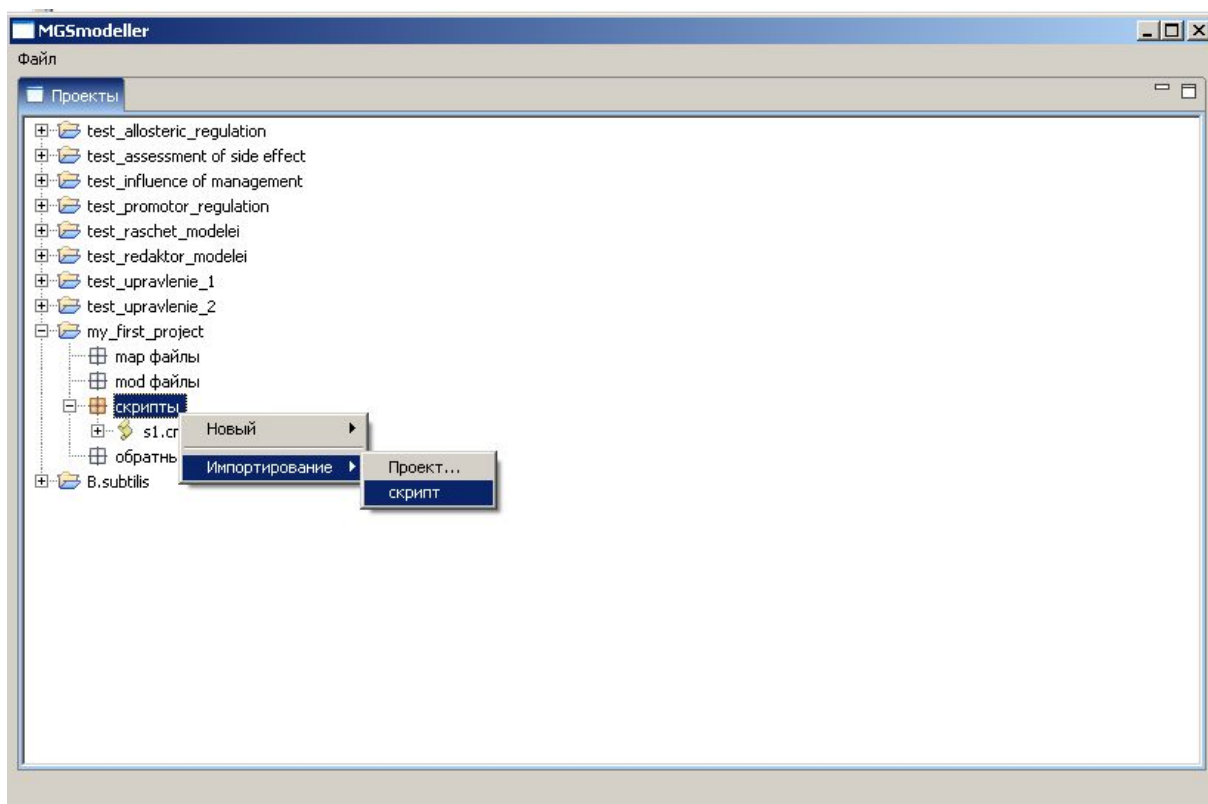
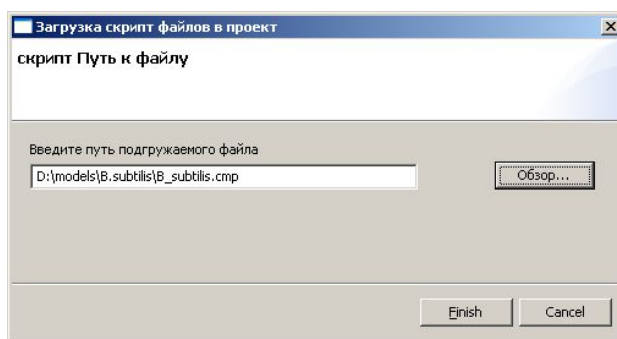


Рисунок 2.1.1.-13 Импортирование скрипта в текущий проект.

В появившемся диалоге введите путь к импортируемому скрипту.



После нажатия кнопки “Finish”. создается копия выбранного файла с расширением «сmp» в текущем проекте. Выполнение данной операции не влечет за собой импортирование mod и map-файлов связанных с данным скриптом.

Создание новых mod и map-файлов в текущем проекте

Для создания нового mod(map)-файла нажмите правой кнопкой мыши на поддерево «mod файлы» («map файлы») в дереве текущего проекта и выберите в выпадающем меню пункт «Новый»/«mod файл» («map файл») (Рис. 14).

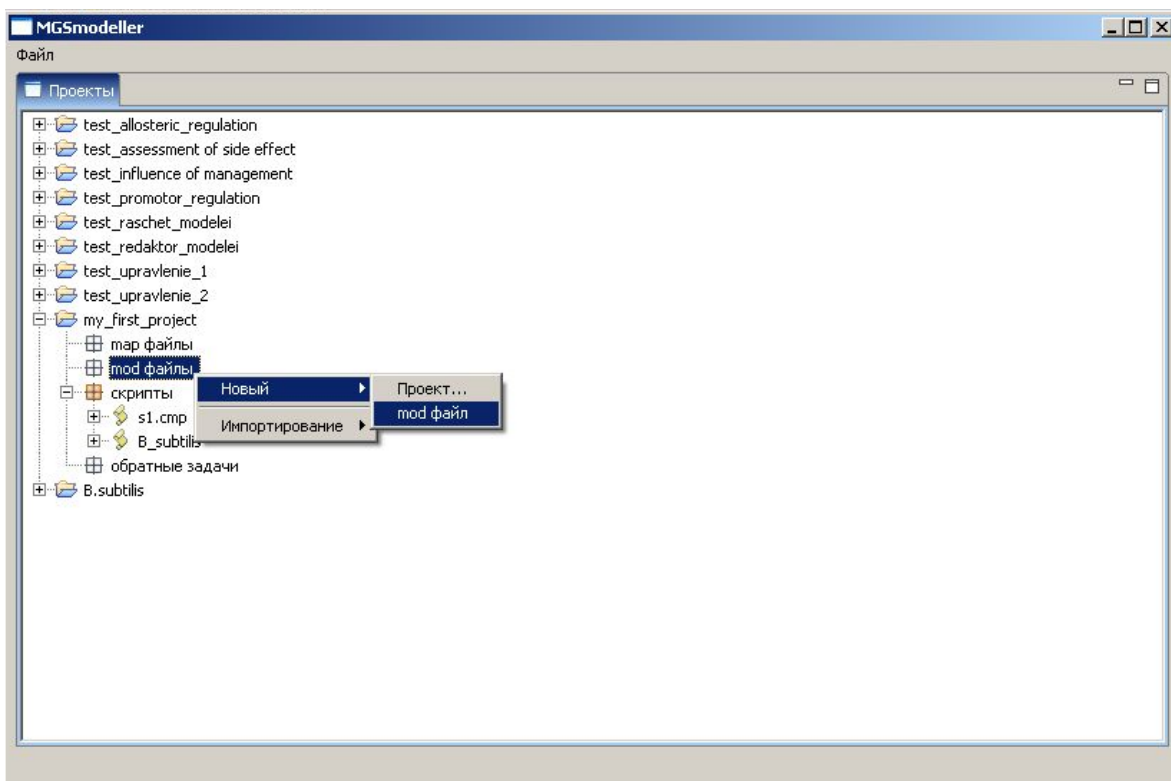
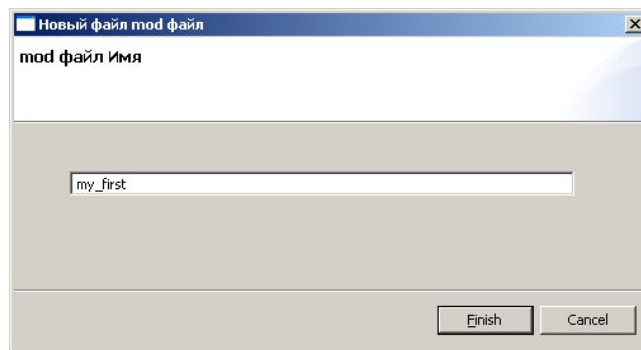


Рисунок 14. Создание нового mod-файла в текущем проекте.

В появившемся диалоге введите имя нового mod(map)-файла.



После нажатия кнопки "Finish" создается новый пустой mod(map)-файл.

Импортирование map-, mod-файлов в текущий проект

Для импортирования нового mod(map)-файла нажмите правой кнопкой мыши на поддерево «mod файлы» («map файлы») в дереве текущего проекта и выберите в выпадающем меню пункт «Импортирование»/«mod файл» («map файл») (Рис. 15).

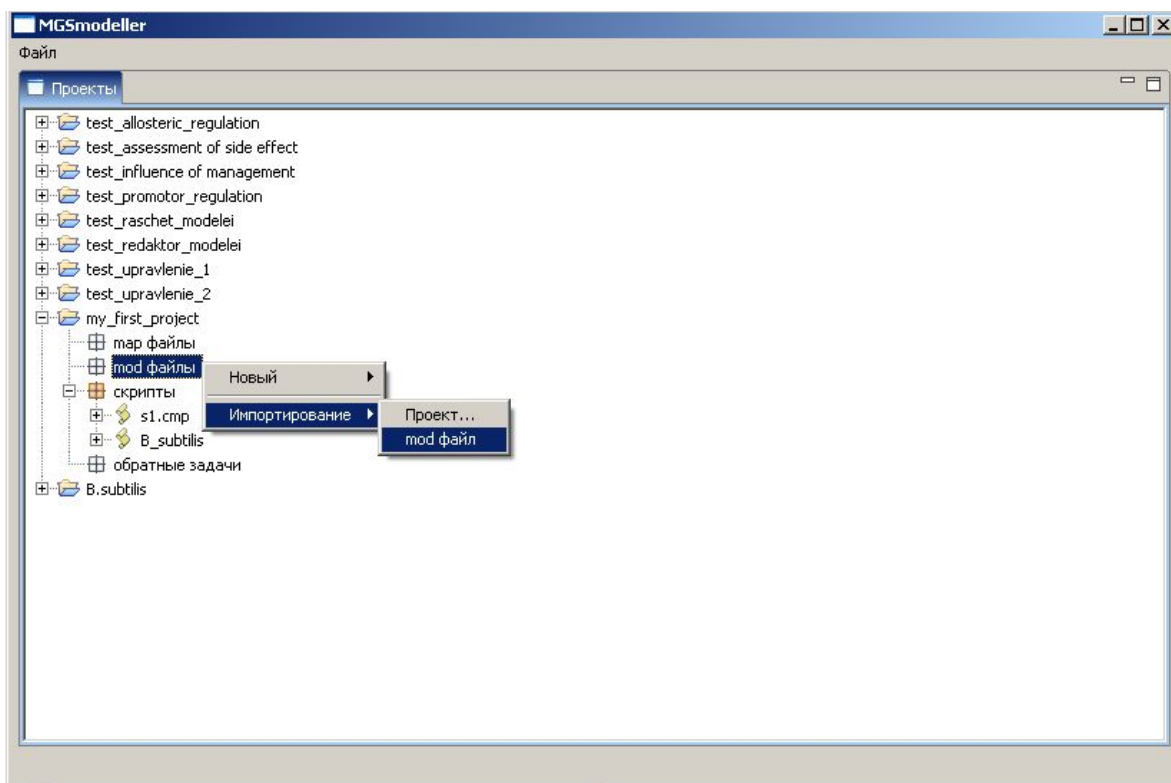
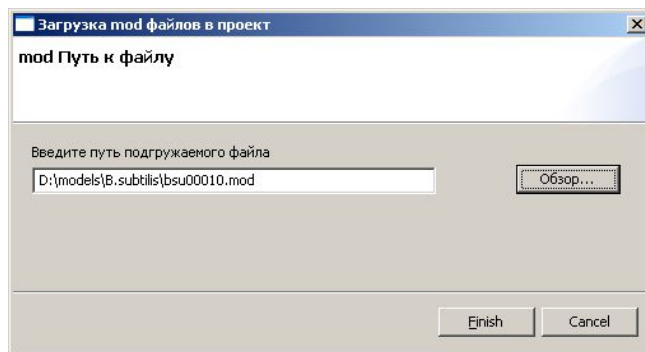


Рисунок 15. Импортирование mod-файла в текущий проект.

В появившемся диалоге введите путь к импортируемому mod(mar)-файлу.



После нажатия кнопки “Finish” mod(mar)-файл импортируется в текущий проект.

2.1.1.3. Постановка и расчет параметрической обратной задачи (ПОЗ)

Постановка обратной задачи

Для постановки обратной задачи выберите проект, в котором хотите поставить обратную задачу. Далее щелкните правой кнопкой мыши на вкладке «Обратные задачи», в появившемся меню выберите пункты «Новый»/«Постановка обратной задачи» (Рис.16).

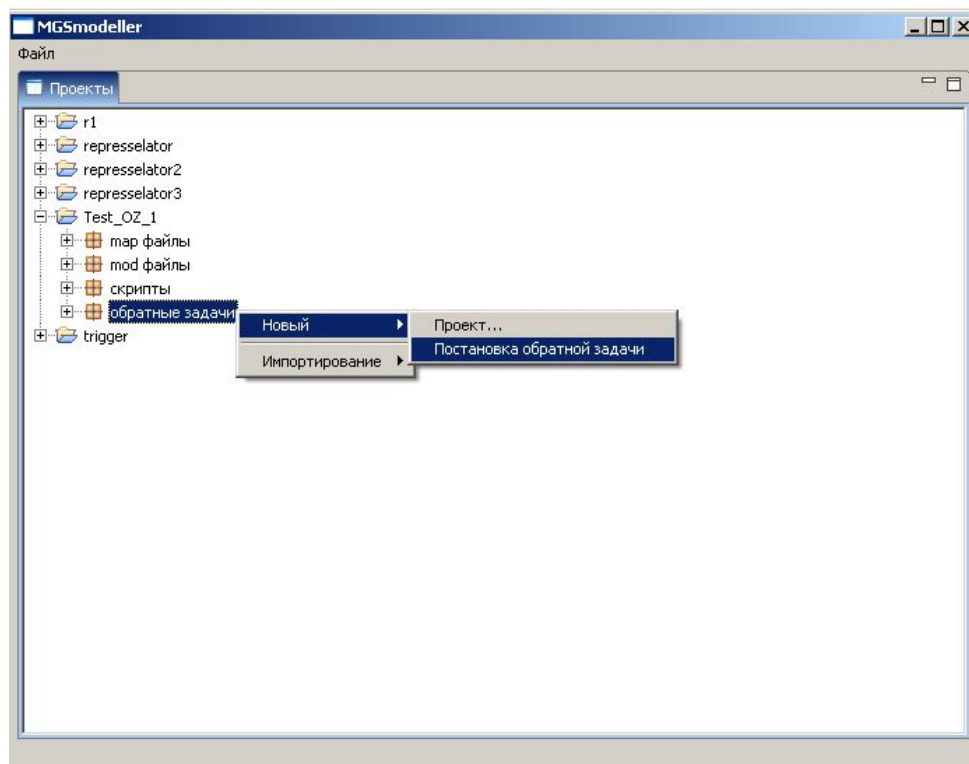


Рисунок 16. Меню «Постановка обратной задачи»

В результате появится мастер постановки обратной задачи. Первая страница мастера содержит доступные скрипты текущего проекта. Необходимо выбрать только один скрипт (двойным нажатием левой клавиши мыши на названии скрипта) (Рис. 17).

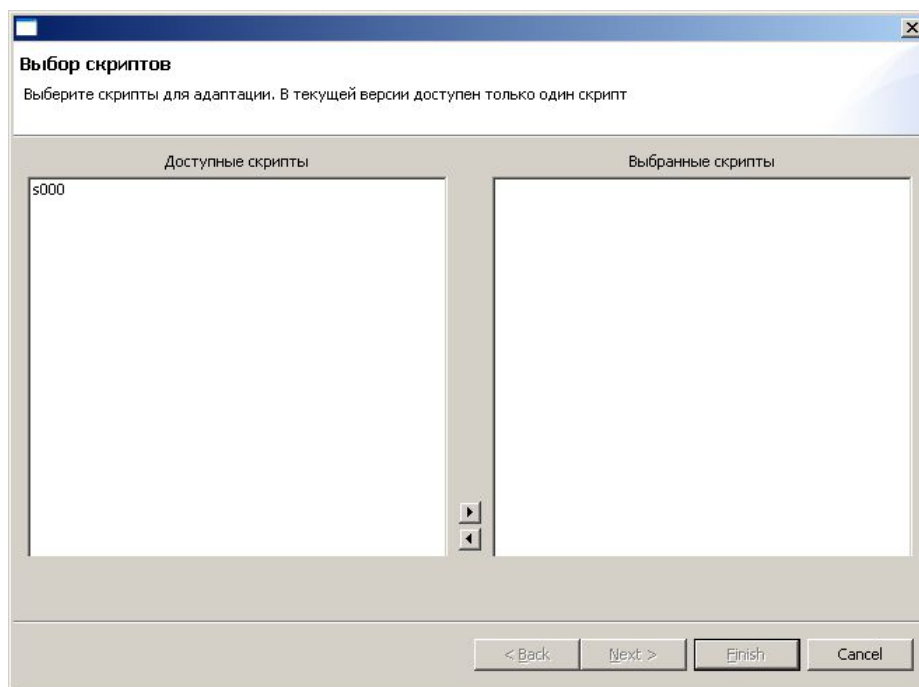


Рисунок 17. Диалоговое окно «Выбор скрипта»

Нажмите кнопку «Далее». На следующей странице мастера отображаются все параметры модели (Рис. 18). Выберите параметры, которым Вы хотите идентифицировать (выбор переменной осуществляется с помощью переключателей (checkbox)). Далее установите вручную начальное (предполагаемое) значение для каждого из выбранных параметров, а также верхнюю и нижнюю границы изменения каждого параметра, либо оставьте значения по умолчанию.

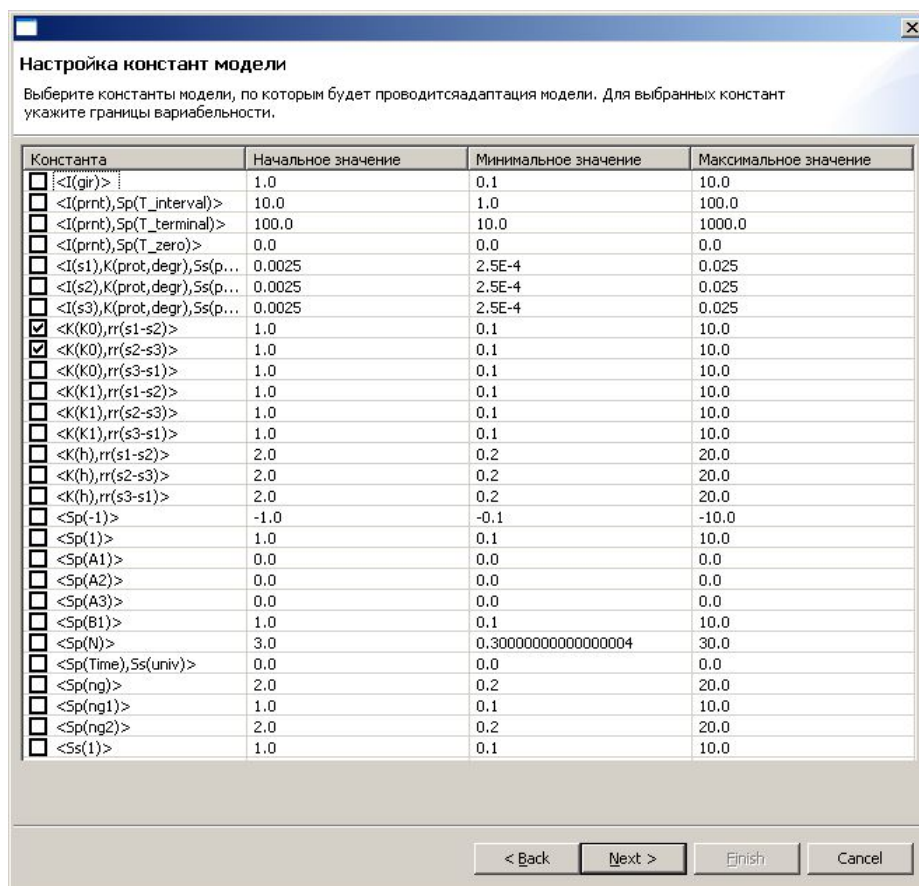


Рисунок 18. Диалоговое окно «Настройка констант модели»

Нажмите кнопку «Далее». На следующей странице мастера (Рис.19) выберите функционал, по которому будет рассчитываться отклонение модельных (расчетных) данных от экспериментальных данных (нажатием левой кнопки на названии метода в выпадающем списке).

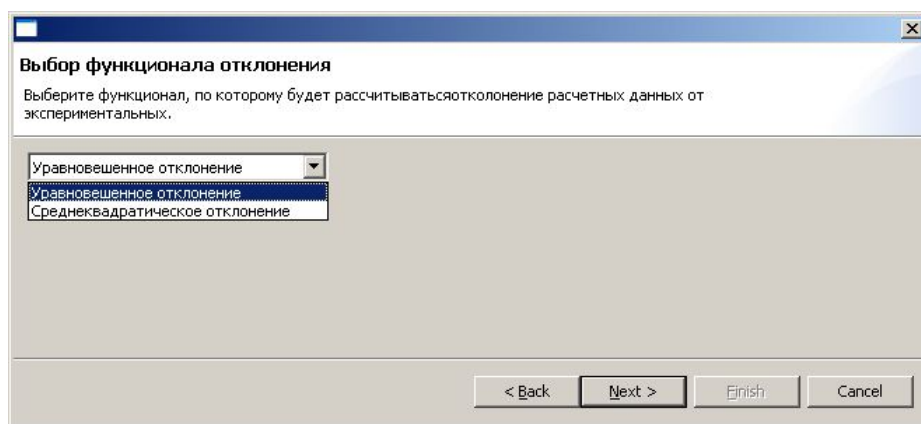


Рисунок 19. Диалоговое окно «Выбор функционала отклонения»

Нажмите кнопку «Далее». На следующей странице мастера (Рис. 20) введите название, под которым конфигурация будет сохранена в проекте. После этого нажмите кнопку «Готово», будет выведено сообщение об успешном создании конфигурации (Рис. 21) и конфигурация будет сохранена в проекте. Также Вы можете нажать кнопку «Отмена» для отмены всех произведенных действий.

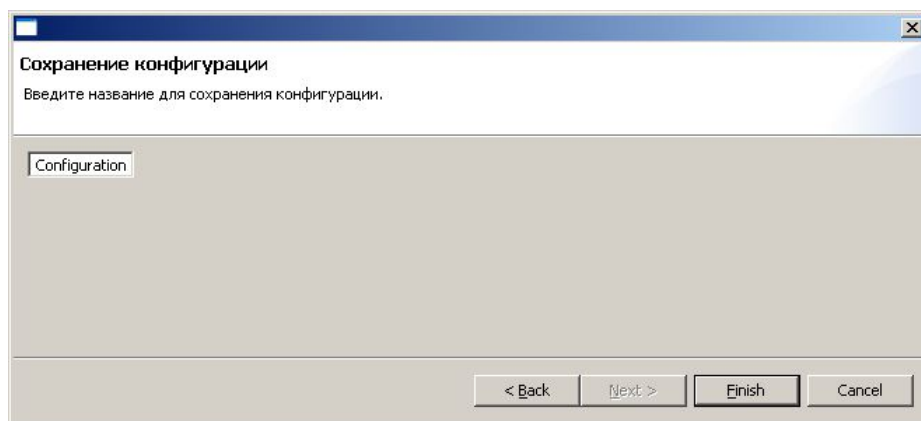


Рисунок 2.1.1.-20 Диалоговое окно «Сохранение конфигурации ПОЗ»

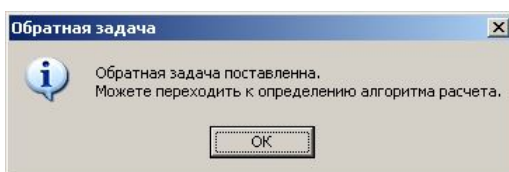


Рисунок 2.1.1.-21 Сообщение об успешном создании конфигурации ПОЗ

Настройка алгоритма расчета обратной задачи

Для настройки алгоритма расчета обратной задачи выберите проект, а в проекте, во вкладке «Обратные задачи», выберите конфигурацию, для которой Вы хотите определить алгоритм расчета (Рис. 22).

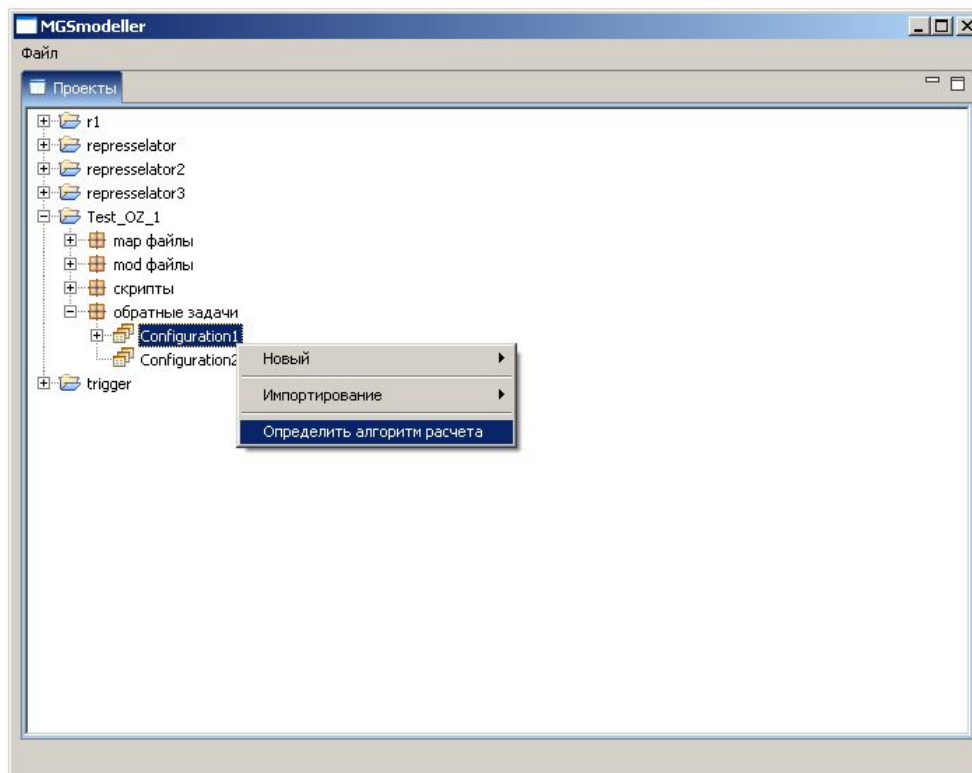


Рисунок 22. Выбор конфигурации ПОЗ, для которой будет определяться алгоритм расчета

В результате появится мастер настройки алгоритма расчета, содержащий список доступных методов решения обратной задачи (Рис. 23). Выберите любой из доступных методов (в текущей версии доступен только Генетический Алгоритм).

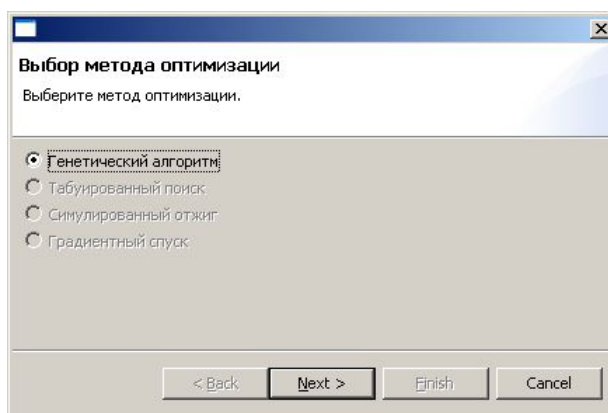


Рисунок 23. Диалоговое окно «Выбор метода решения ПОЗ»

Нажмите кнопку «Далее». Следующая страница мастера содержит настройки выбранного алгоритма. Установите параметры алгоритма вручную, либо нажмите кнопку «По умолчанию», которая установит параметры автоматически (Рис. 24).

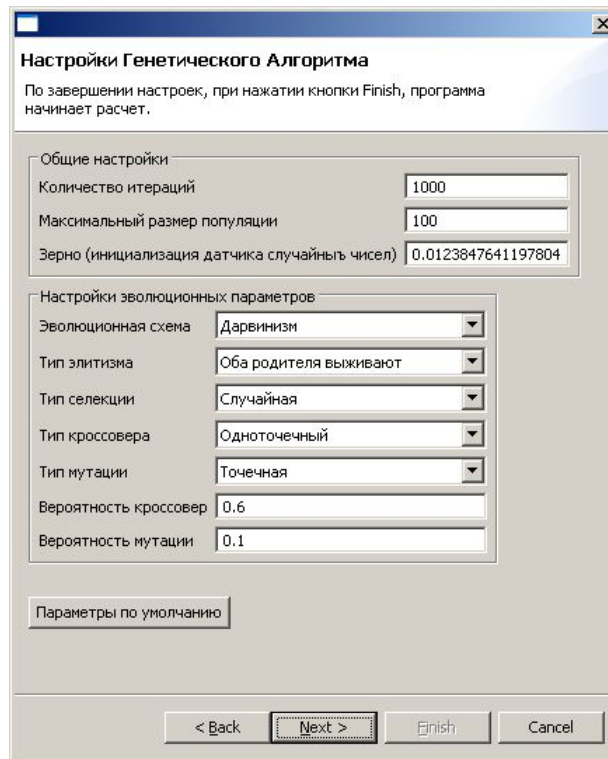


Рисунок 24. Диалоговое окно «Настройка параметров генетического алгоритма»

Нажмите кнопку «Далее». На следующей странице мастера введите название, под которым настройки алгоритма расчета сохранятся в проекте (Рис. 25). После этого нажмите кнопку «Готово», настройки будут сохранены в проекте. Также Вы можете нажать кнопку «Отмена» для отмены всех произведенных действий.

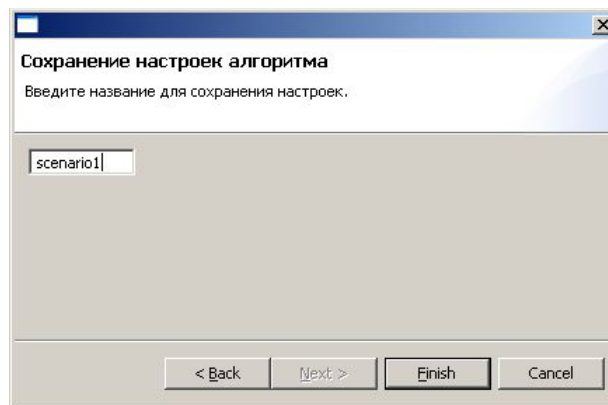


Рисунок 25. Диалоговое окно «Сохранение настроек алгоритма»

Расчет обратной задачи

Для расчета обратной задачи необходимо, чтобы скрипт, для которого будет рассчитываться обратная задача, содержал файл экспериментальных данных – «эталонный файл» (подробнее о составлении эталонных файлов см. Руководство пользователя), а также,

чтобы в скрипте был outdat-файл, оформленный таким образом, чтобы выборочный вывод при расчете соответствовал структуре эталонного файла (также см. Руководство пользователя, там же). Кроме этого, необходимо, чтобы проект содержал постановку обратной задачи, а также, чтобы постановка содержала хотя бы один настроенный метод решения ПОЗ. Выберите в текущем проекте вкладку «Обратные задачи», выберите конфигурацию и метод расчета, правой кнопкой мыши вызовите контекстное меню и выберите пункт «Запустить на счет» (Рис. 26).

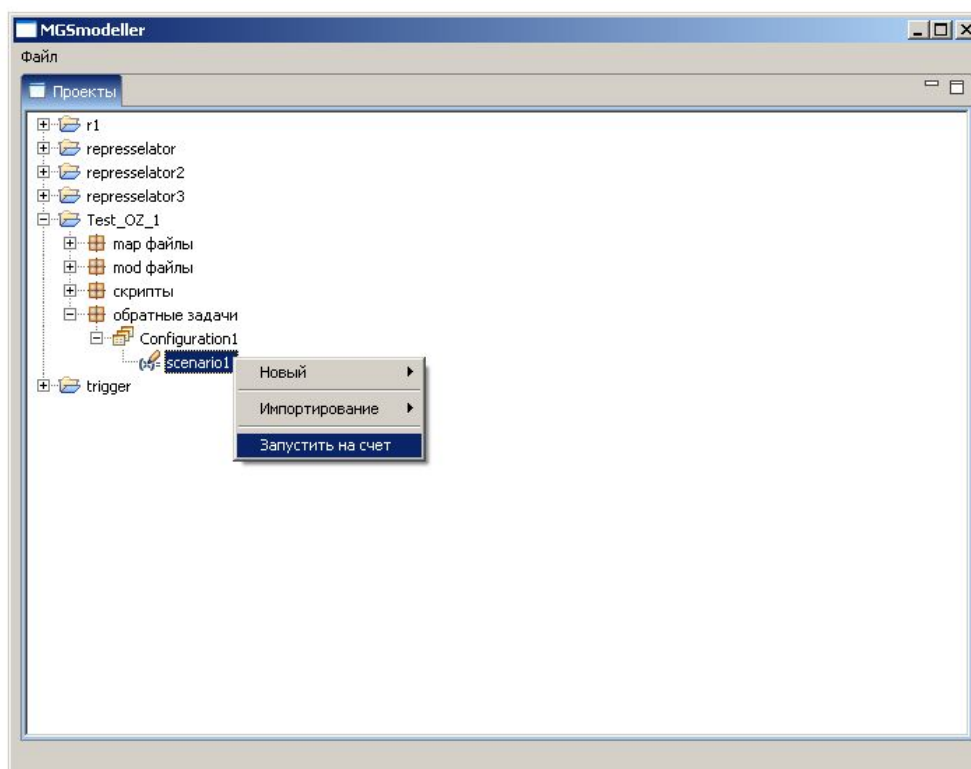


Рисунок 26. Запуск решателя обратной задачи

После решения обратной задачи появится сообщение об успешном завершении счета, которое будет содержать достигнутое оптимальное значение функционала отклонения (Рис. 27).

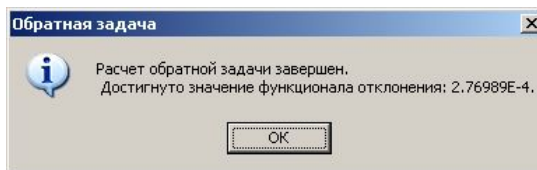


Рисунок 27. Сообщение о завершении расчета ПОЗ

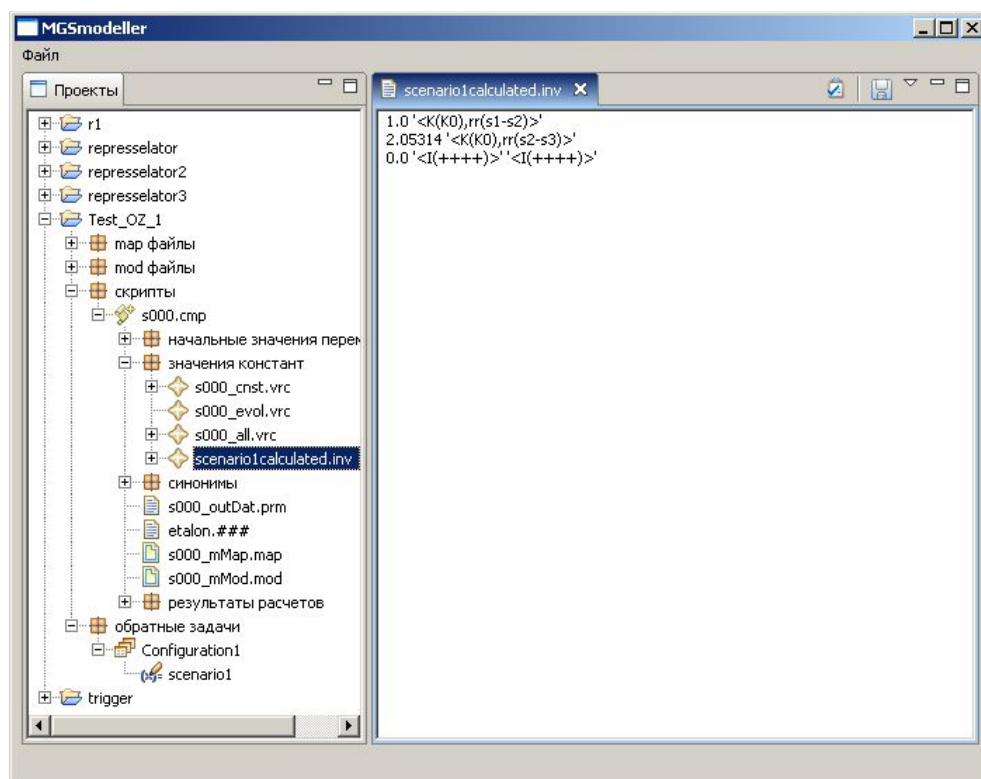


Рисунок 28. Пример файла с идентифицированными константами.

Во вкладке скрипта «значения констант» появится файл «*calculated.inv» (где символом «*» обозначено название конфигурации обратной задачи, для которой проводился расчет). Этот файл содержит значения констант, которые были идентифицированы (Рис. 28).

Постановка и решение задачи оптимального управления

Задача оптимального управления может быть поставлена для любого собранного скрипта.

Условие задачи хранится в управляющем блоке «UPRV» в файле “*_upravlenie.mod” (где символом “*” обозначено имя скрипта) и подключается к скрипту через сmp-файла (строка, определяющая задачу оптимального управления выделена серым) с помощью мар-файл “*_upravlenie.mar” (где символом “*” обозначено имя скрипта):

```
'<ACTCOMP#(gear) MAP_ADRES#%(gear.map) MOD_ADRES#%(gear_block.mod) stroka(+)>'
'<ACTCOMP#(prnt) MAP_ADRES#%(prnt.map) MOD_ADRES#%(tprm_block.mod) stroka(+)>'
'<ACTCOMP#(upravlenie) MAP_ADRES#%(s000_upravlenie.map) MOD_ADRES#%(s000_upravlenie.mod) stroka(+)>'
'<ACTCOMP#(0) MAP_ADRES#%(main.map) MOD_ADRES#%(main.mod) stroka(+)>'
'<C(++++)>'
```

Таким образом, условие задачи управления является частью скрипта, может быть изменено на любом шаге решения задачи управления. Любое изменение в условии задачи управления требует пересборки скрипта. Изменение условия задачи управления или постановка новой задачи может быть проведена напрямую через блок “UPRV”. Постановка новой задачи управления проводится с помощью мастера «Постановка задачи управления».

Постановка задачи управления с помощью мастера

Вызовите мастер «постановки задачи управления». Для этого в контекстном меню скрипта выберите пункт «Задача управления»/«Постановка задачи» (Рис. 29).

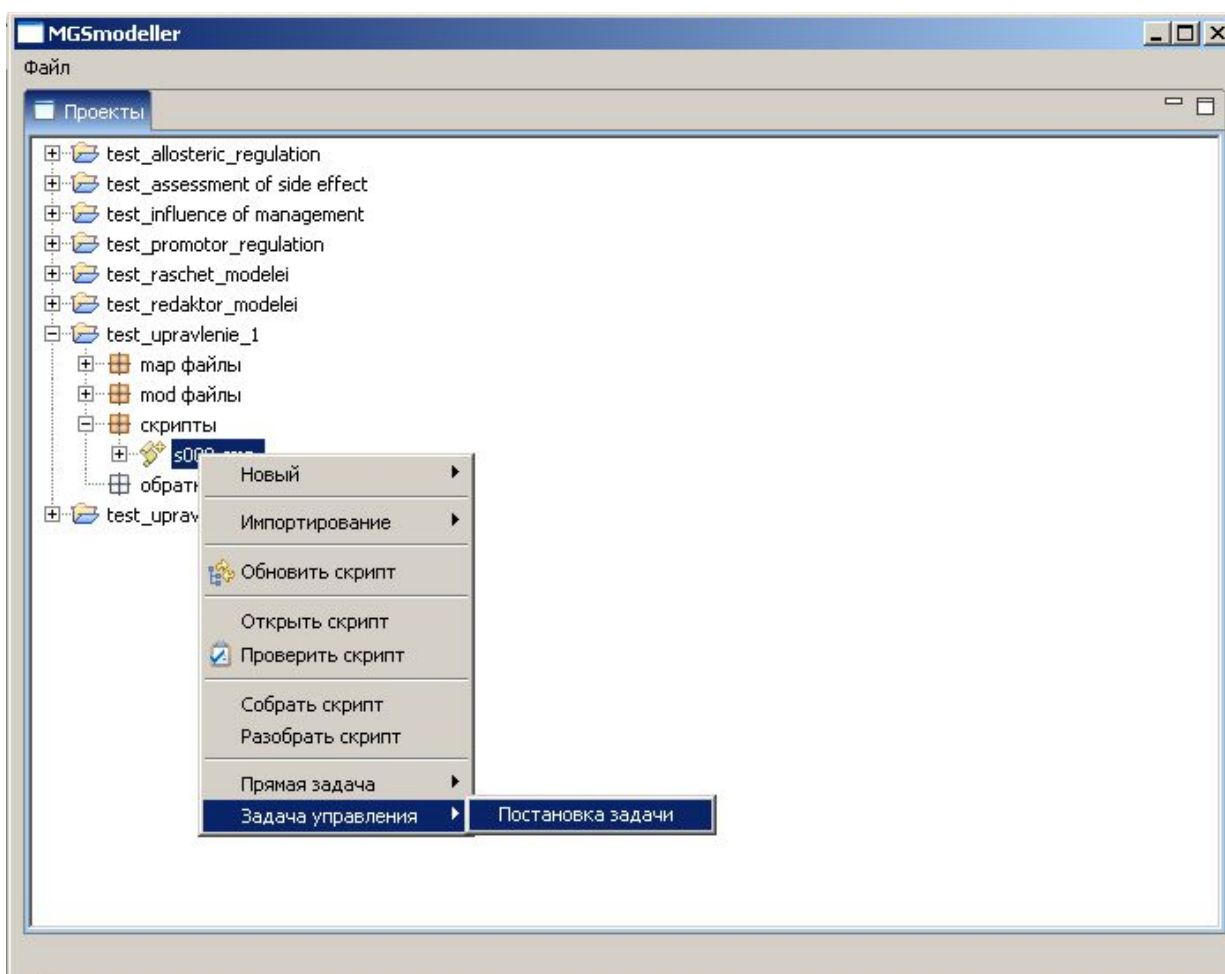


Рисунок 29. Запуск мастера постановки задачи управления.

Откроется диалоговое окно «Задание управляющих переменных» (Рис. 30). Выберите динамические переменные, которые будут участвовать в управлении.

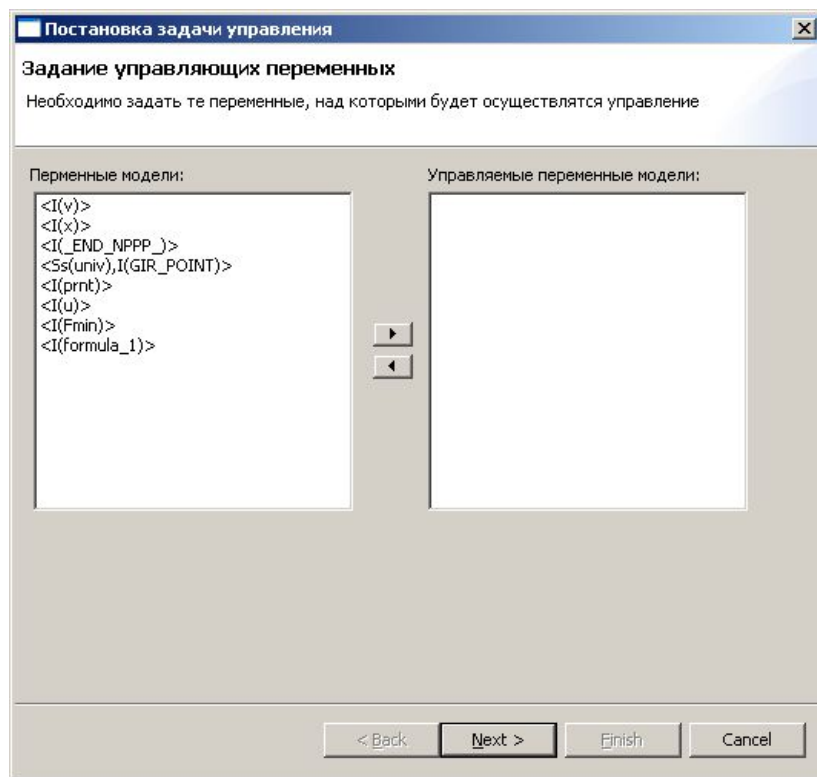


Рисунок 30. Диалоговое окно «Задание управляющих переменных»

Окончание процесса выбора динамических переменных соответствует нажатию левой клавишей мыши на кнопку «Далее». После этого возникает окно следующего вида (Рис. 31):

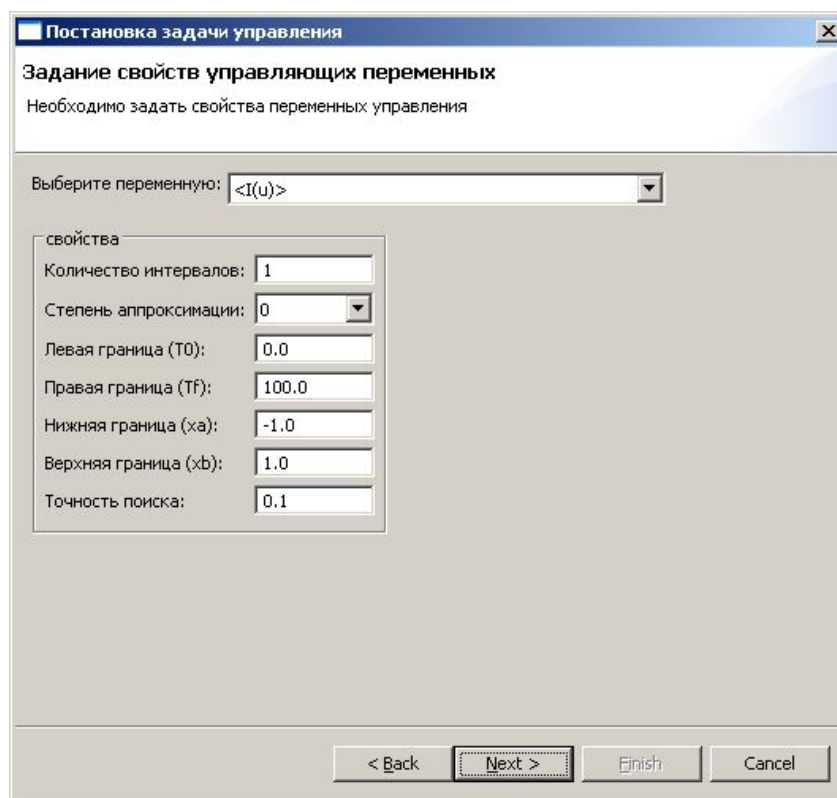


Рисунок 31. Диалоговое окно «Выбор свойств управляющих переменных»

В данном окне специфицируются свойства управляющих переменных модели: количество интервалов, степень аппроксимации, левая и правая границы, нижняя и верхняя границы, а также точность поиска. Для того чтобы задавать свойства управляющих переменных, необходимо левой кнопкой мыши нажать на поле ввода напротив имени свойства и с помощью клавиатуры внести требуемое значение. Окончание процесса задания свойств управляющих переменных модели соответствует нажатию левой кнопкой мыши на кнопку «Далее». После спецификации свойств управляющих переменных возникает окно, в котором пользователь может задавать управляющие переменные из ряда констант модели (Рис. 32):

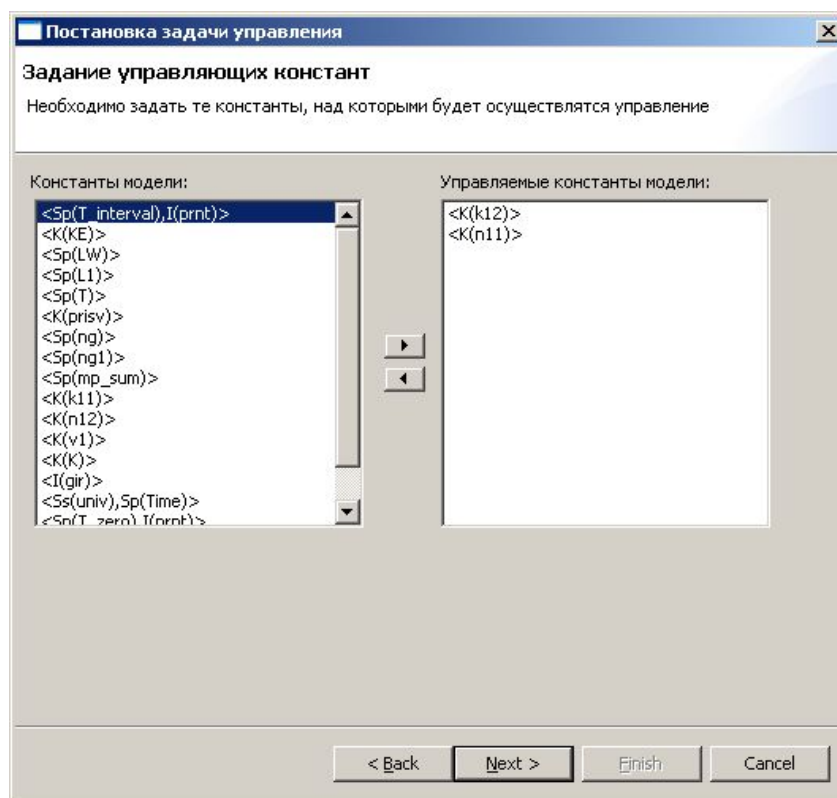


Рисунок 32. Диалоговое окно «Задание управляющих констант»

Окончание процесса выбора управляющих констант соответствует нажатию левой кнопкой мыши на кнопку «Далее». После этого возникает окно следующего вида (Рис. 33):

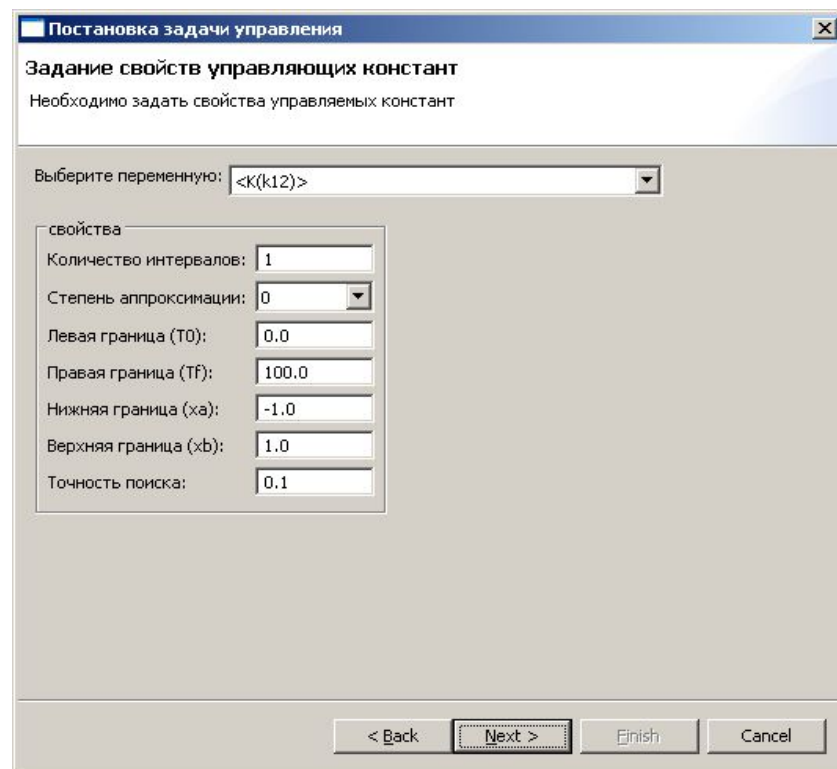


Рисунок 33. Диалоговое окно «Задание свойств управляющих констант»

В данном окне специфицируются свойства управляющих констант модели: количество интервалов, степень аппроксимации, левая и правая границы, нижняя и верхняя границы, а также точность поиска. Для того чтобы задавать свойства управляющих констант, необходимо левой кнопкой мыши нажать на поле ввода напротив имени свойства и с помощью клавиатуры внести требуемое значение. Окончание процесса задания свойств управляющих констант модели соответствует нажатию левой кнопкой мыши на кнопку «Далее». После спецификации свойств управляющих констант возникает окно, в котором пользователь может задавать переменные-ограничения (Рис. 34):

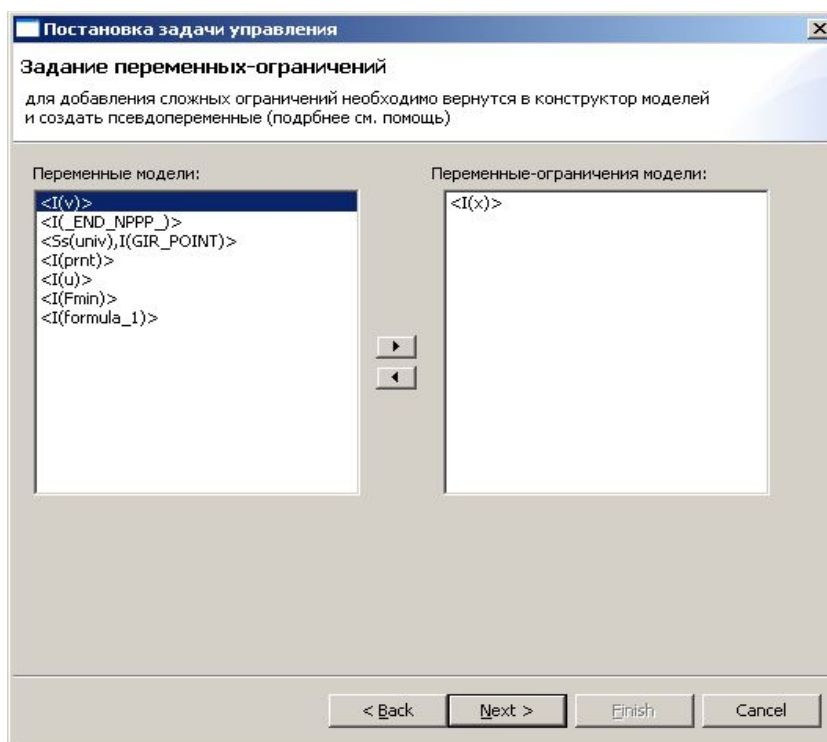




Рисунок 34. Диалоговое окно «Задание переменных-ограничений»

Для того чтобы задавать переменные-ограничения, пользователю необходимо с помощью нажатия левой кнопкой мыши на кнопку  перенести переменные из общего списка переменных модели в список переменных-ограничений. Если пользователь ошибочно перенёс имя переменной в список переменных-ограничений, то нажатием левой кнопки мыши на кнопку , он может вернуть эту переменную в общий список. Окончание процесса задания переменных-ограничений модели соответствует нажатию левой кнопкой мыши на кнопку «Далее».

После этого возникает окно следующего вида (Рис. 35):

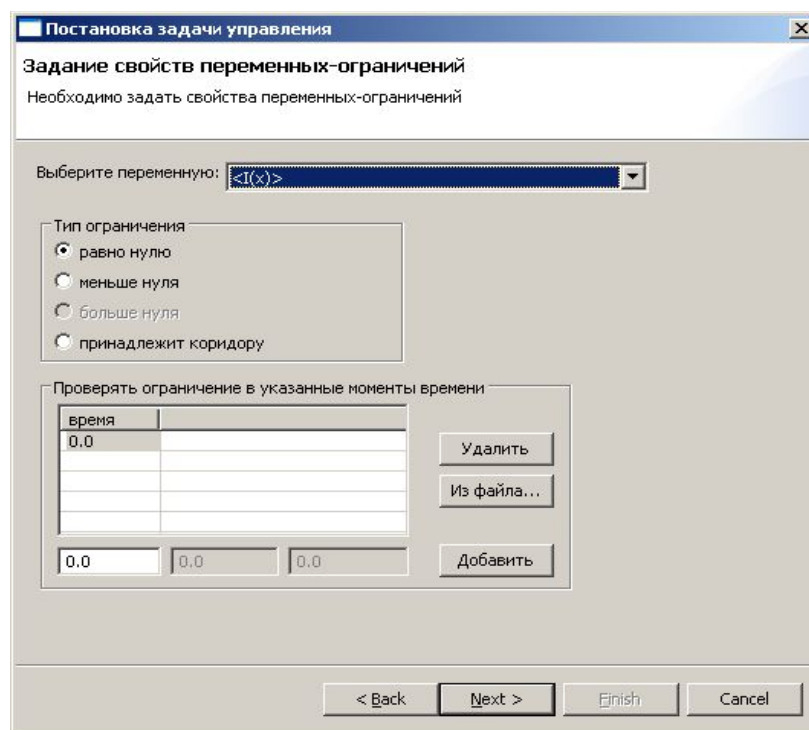


Рисунок 35. Диалоговое окно «Задание свойств переменных-ограничений»

В данном окне специфицируются свойства переменных-ограничений модели: тип ограничений, проверка ограничений в указанные промежутки времени. Для того чтобы выбрать тип ограничений, необходимо левой кнопкой мыши установить флажок напротив типа ограничения. Для того чтобы задавать проверку ограничений в указанные моменты времени, необходимо добавить соответствующие записи в таблицу «Проверять ограничения в указанные моменты времени». Добавление и удаление текущей (выбранной) записи происходит соответственно по нажатию на кнопки «Добавить» и «Удалить». Для того чтобы загрузить список времен и значений левой и правой границы из файла, необходимо нажать на кнопку «Из файла...» и в появившемся диалоге выбрать нужный файл (файл должен содержать значения разделенные пробелом, тубляцией или переводом строки).

Окончание спецификации свойств переменных-ограничений модели соответствует нажатию левой кнопкой мыши на кнопку «Далее». После этого возникает окно следующего вида (Рис. 36):

Рисунок 37. Диалоговое окно «Задание свойств переменной-функционала»

В данном окне задаете моменты времени, в которые необходимо минимизировать выбранный функционал (переменную). Результирующий функционал будет определяться как сумма значений выбранного функционала в указанные моменты времени. Для того чтобы загрузить список времен из файла, необходимо нажать на кнопку «Из файла...» и в появившемся диалоге выбрать нужный файл (файл должен содержать только моменты времени разделенные пробелом, табуляцией или переводом строки). Редактирование таблицы осуществляется с помощью кнопок «Добавить» и «Удалить». Окончание задания моментов времени для переменной-функционала соответствует нажатию левой кнопкой мыши на кнопку «Далее». После этого возникает окно следующего вида (Рис. 37):

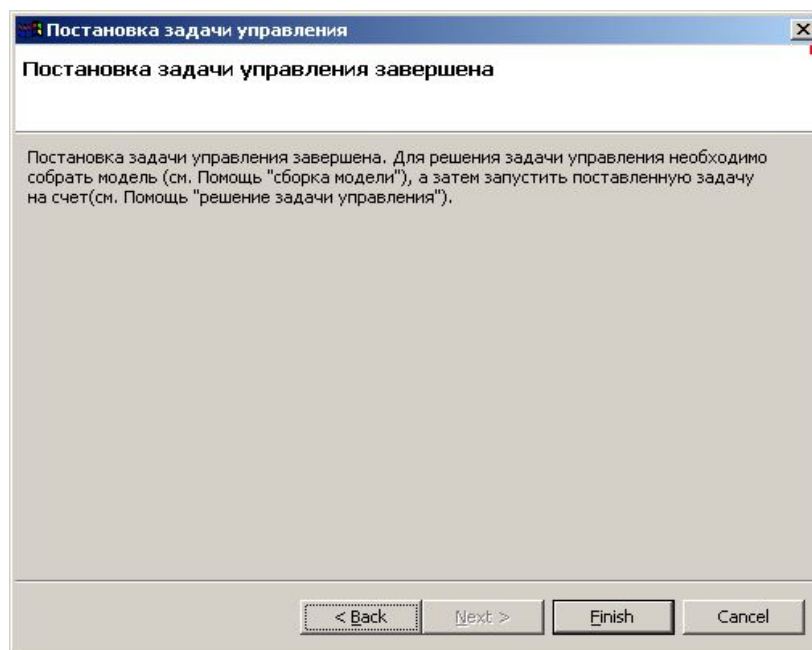


Рисунок 38. Диалоговое окно «Постановка задачи управления завершена»

Для завершения постановки задачи управления необходимо нажать на кнопку «Закончить».

Как только вы закончили работу с мастером по нажатию кнопки «Закончить», в проект добавились файлы “*_upravlenie.mod” и “*_upravlenie.map” (где символом “*” обозначено имя скрипта), в стр-файл скрипта добавилась строчка:

```
'<ACTCOMP#(upravlenie) MAP_ADRES#%(s000_upravlenie.map) MOD_ADRES#%(s000_upravlenie.mod) stroka(+)>'
```

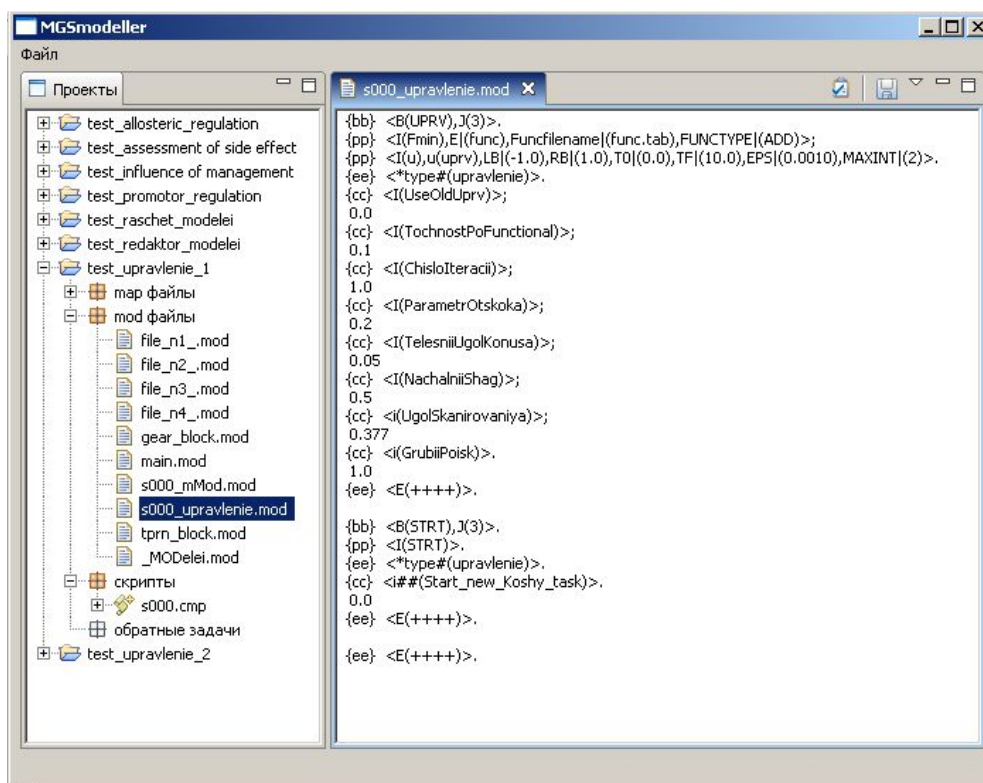


Рисунок 39. Пример файла с постановкой задачи управления.

Файл с постановкой задачи управления “*_upravlenie.mod” (где символом “*” обозначено имя скрипта) добавляется к текущему проекту (Рис. 39).

Запуск на счет задачи управления

Запуск задачи управления осуществляется следующим образом. Соберите скрипт с постановкой задачи управления (см. раздел «Сборка скрипта»). Запустите на счет задачу управления выбрав в контекстом меню скрипта пункт «Задача управления»/«Запустить на счет» (Рис. 40).

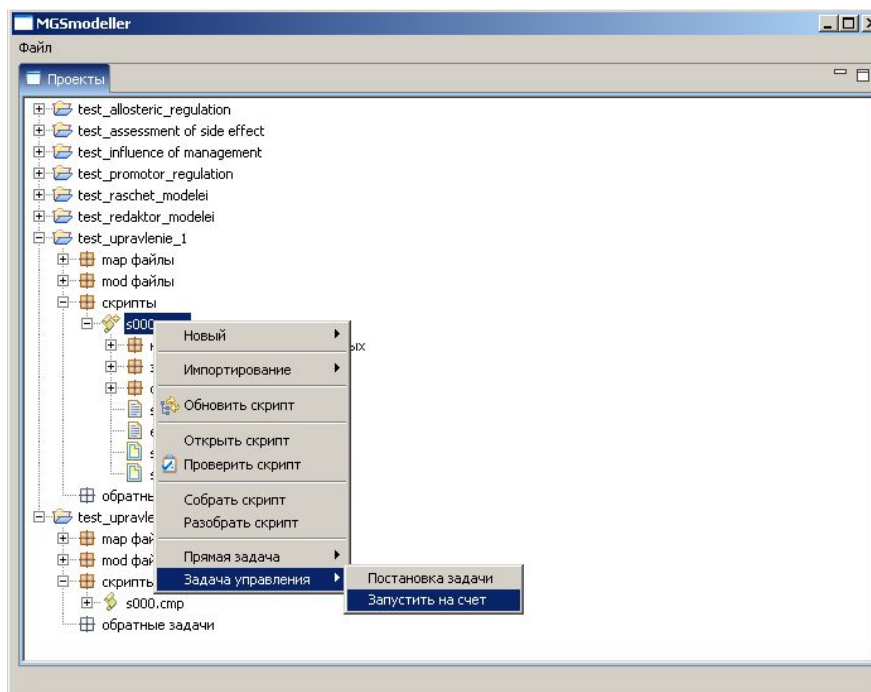
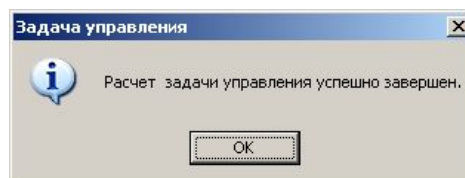


Рисунок 40. Запуск на счет задачи управления.

После окончания решения задачи, выводится следующее окно:



Результаты решения задачи управления пользователь может посмотреть в файлах «*_evol.vpr» и «*_evol.vtc» (где символом «*» обозначено имя скрипта), в которых отображаются соответственно динамика управляющих переменных и констант модели (Рис. 41):

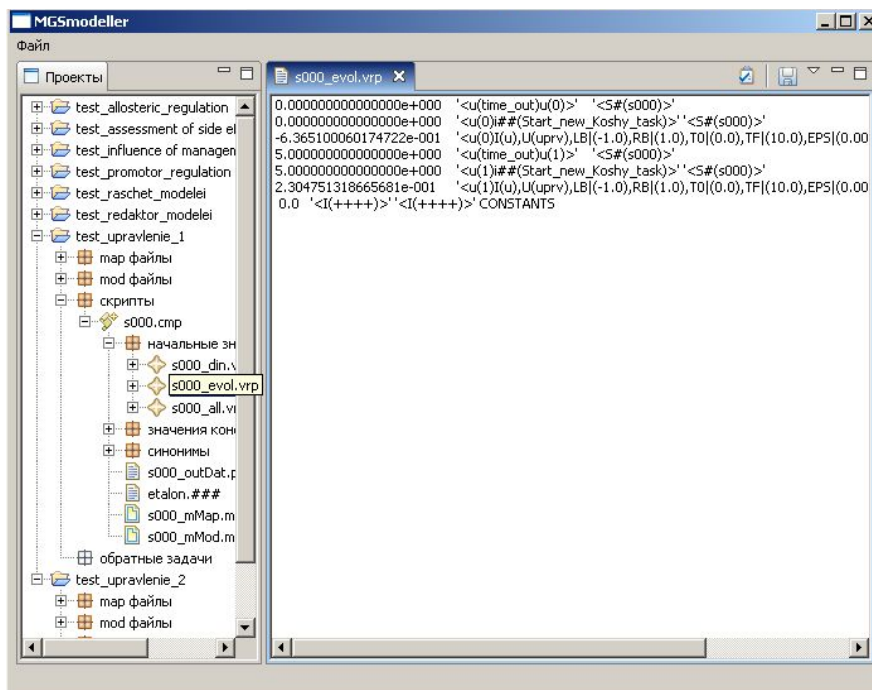
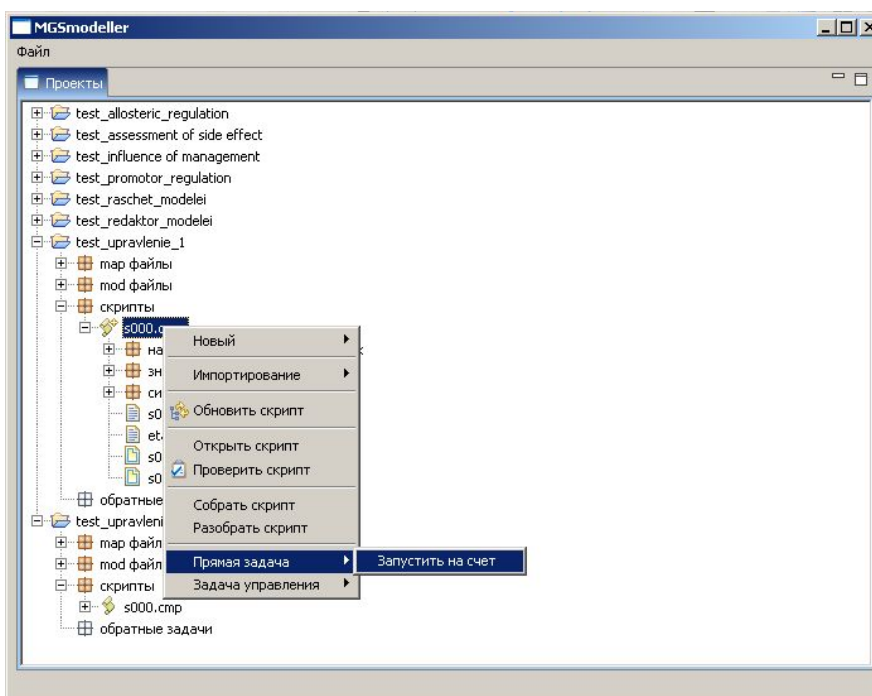


Рисунок 41. Результат решения задачи управления.

Расчет прямой задачи с управляющими воздействиями

Расчет прямой задачи с управляющими воздействиями, полученными после решения задачи управления или с любыми другими прописанными в файлах “*_evol.vrp” и “*_evol.vrc” (где символом “*” обозначено имя скрипта) (о формате записи управляющих воздействий см. подробнее в Руководстве пользователя) осуществляется выбором пункта контекстного меню скрипта «Прямая задача»/«Запустить на счет» (Рис. 42):



2.1.2. Программная компонента «METABOL»

2.1.2.1. Структура и функциональные возможности программного компонента «METABOL».

Программный компонент "METABOL" предназначен для моделирования и анализа полных (включая обмен со средой) и частичных систем метаболических реакций у бактерий на основе линейного программирования. Программный компонент обеспечивает нахождение установившихся величин потоков для заданной системы метаболических реакций, целевой функции и ограничений на потоки.

«METABOL» моделирует метаболизм бактерии в стационарном состоянии, включая химический состав среды, из которой бактерия получает питание. Это позволяет, варьируя состав среды, наблюдать, какие метаболические пути функционирования бактерии в рамках изучаемой модели активируются, а какие выключаются. И, тем самым, находить альтернативные метаболические пути для обеспечения жизнедеятельности клетки, для обеспечения выхода биомассы выше достигнутого уровня. «METABOL» позволяет по каждой реакции задать количество соответствующего фермента как верхнюю и нижнюю границы на значение интенсивности реакции. Это позволяет направлять биохимию клетки по определенным путям биосинтеза и изучать зависимость биохимии от величины экспрессии соответствующего гена. Варьируя состав среды и количества ферментов, можно искать режимы работы клетки, при которых происходят интересные исследователя процессы, например, наиболее интенсивно вырабатываются кофакторы или биомасса достигает максимального значений. Таким образом, можно рассматривать как живые, так и модельные организмы. «METABOL» также позволяет анализировать устойчивость метаболизма к изменениям величин, характеризующих среду и экспрессию генов.

Применение программы предполагает наличие подготовленного в специальном формате MPS (позиционный формат IBM для описания задач линейного программирования) представления метаболической системы интересующей бактерии в виде совокупности метаболических реакций и участвующих в них веществ, а также характеристик внешней среды и дополнительных ограничений, характеризующих конкретную решаемую задачу. Учитывая большой объем и разнородность этих данных, для построения файлов в формате MPS предлагается технология, в соответствии с которой указанная информация вначале собирается и редактируется в электронной таблице, а затем с помощью отдельной программы переносится в формат MPS, одновременно проходя контроль безошибочности.

Понимание текста размещенного в разделе 2.1.2. подразумевает, что пользователь ознакомился со стандартом и форматом файлов mps, применяемых в программном компоненте "METABOL". Подробное описание стандарта и формата данного файла дано в "ОИОС01". Подробное описание руководства пользователя дано в "АСНИ-01".

2.2.2.2. Создание компьютерных моделей систем метаболических реакций у бактерий в программной компоненте «METABOL».

Запуск METABOL

Запустите исполняемый файл metabol_win.exe в директории, в которой размещен программный компонент «МЕТАВОЛ».
Появляется диалоговое окно выбора файлов модели (Рис. 43).

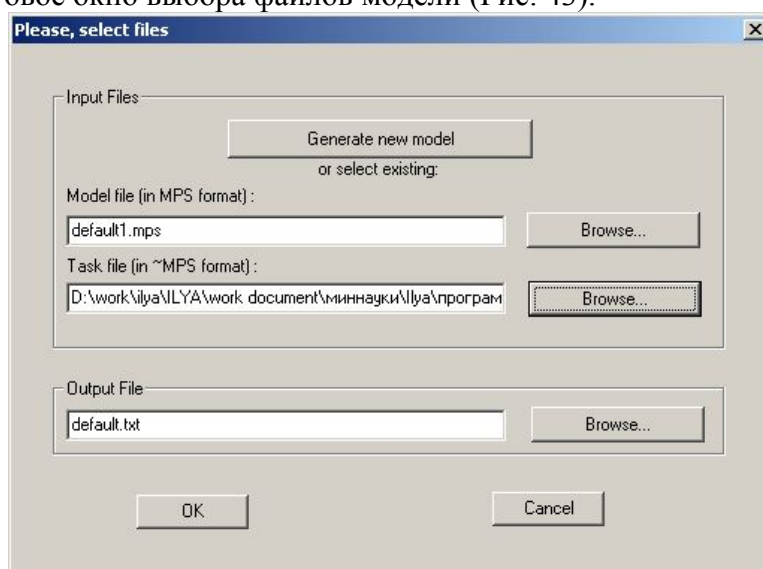


Рисунок 43. Диалоговое окно выбора файлов модели

Создание новой модели

Нажмите кнопку “Generate new model” (создать новую модель) в диалоговом окне выбора файлов модели (Рис. 43).

Появляется диалоговое окно создания новой модели (Рис. 44).

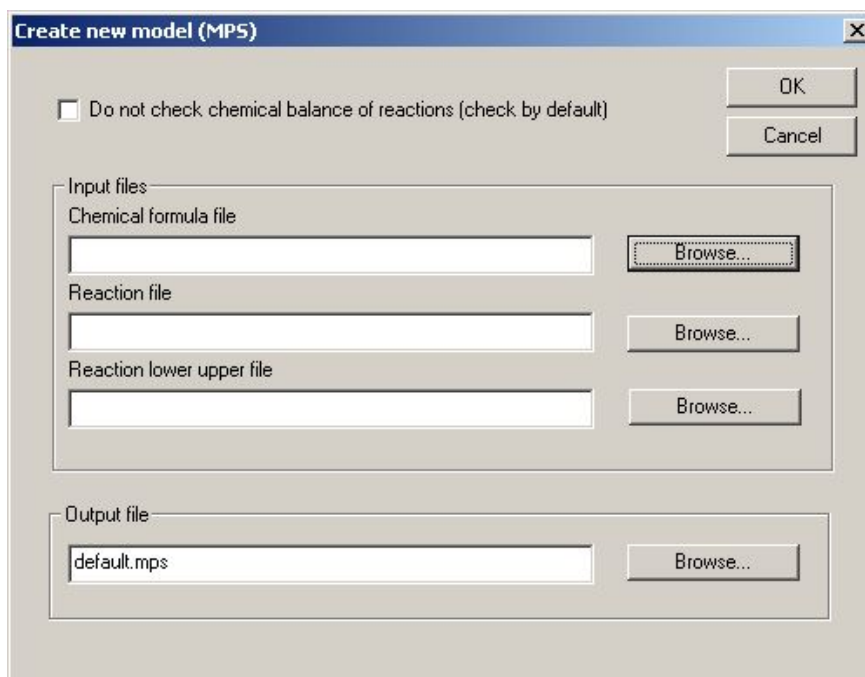


Рисунок 44. Диалоговое окно создания новой модели

Выбор файлов для создания новой модели осуществляется в результате выбора директории, в которой хранится файл.

Ввод пути к файлам, где находятся созданные входные данные модели, можно осуществлять вручную, либо используя вспомогательный диалог, который активизируется после нажатия кнопки “Обзор” (Browse) (Рис. 45).

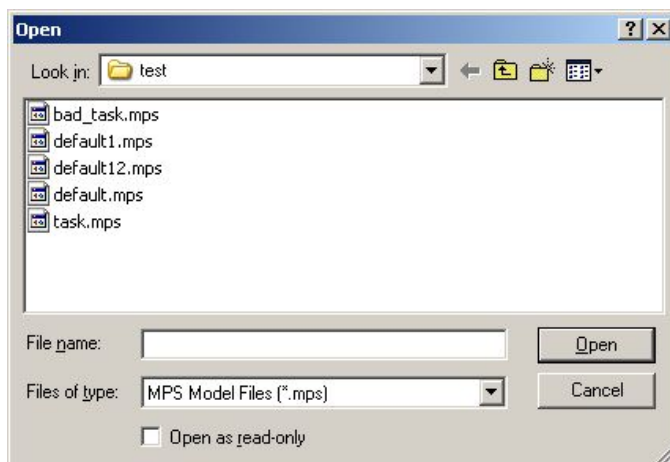


Рисунок 45. Диалоговое окно выбора директории входных файлов

Для загрузки файлов, содержащих обучающие примеры, выберите поддиректорию test в директории, в которой размещен «МЕТАБОЛ». Выберите последовательно файлы для создания новой модели: abbr, react, exch.

Введите имя файла новой модели default1.mps в поле ввода имени файла mps (Рис. 44). Нажмите ОК. Новая модель успешно создана.

Загрузка существующей модели

Выберите файлы default.mps и task.mps в поддиректории test (Рис. 45).

Введите имя выходного файла модели default.txt в поле ввода имени выходного файла (Output file) (Рис. 43).

Сборка модели и расчёт прямой задачи

Нажмите ОК в диалоговом окне выбора файлов модели.

Появляется диалоговое окно опций программы (Рис. 46). Заполните опции этого диалогового окна установкой флажков напротив имени опции.

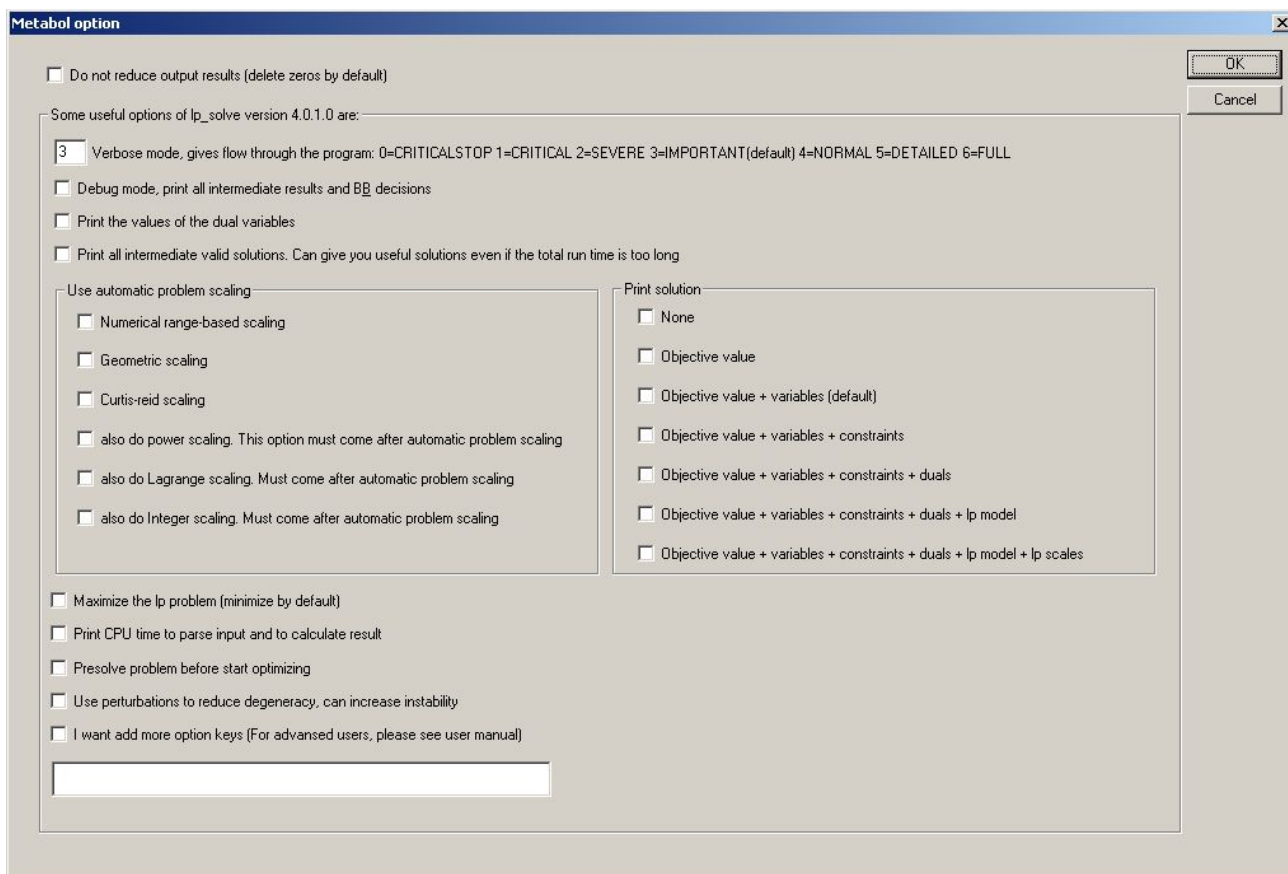


Рисунок 46. Диалоговое окно опций программы

Нажмите ОК. Новая модель успешно создана и рассчитана.

После сборки и расчёта модели появляется рабочее пространство программы (Рис. 47).

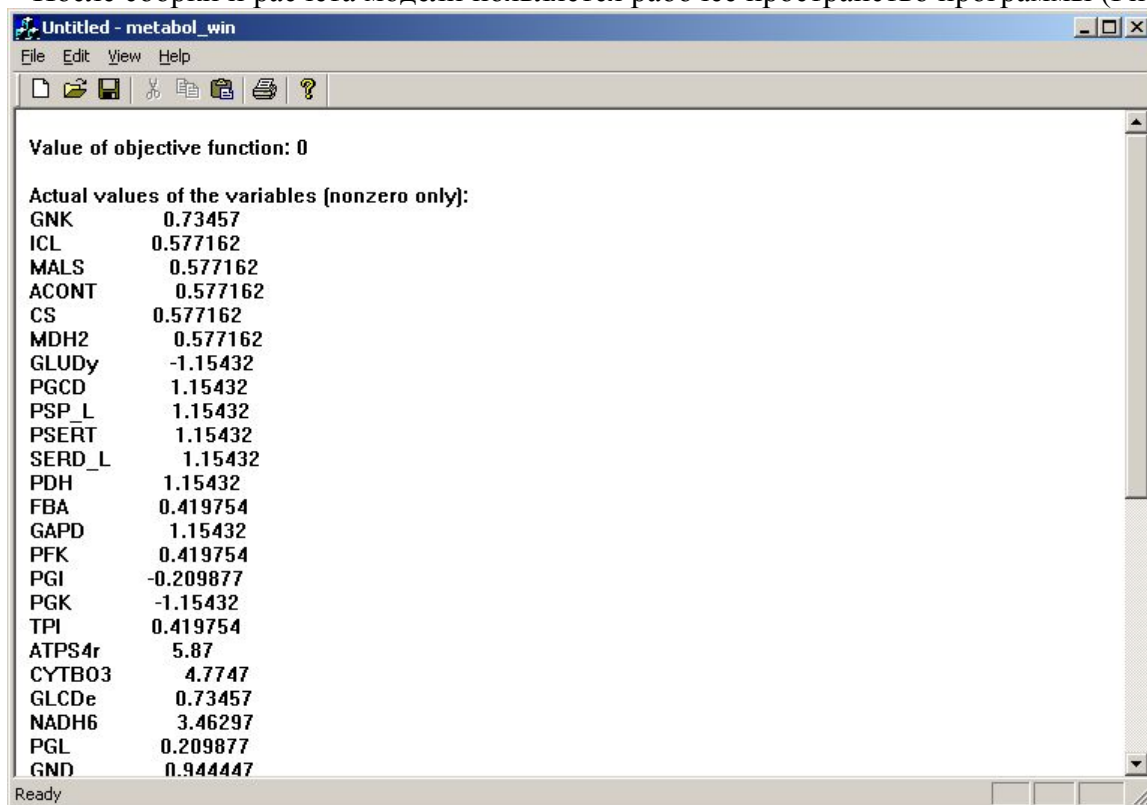


Рисунок 47. Рабочее пространство программы

Для редактирования соответствующего файла программы откройте его через всплывающее меню View в Главном меню программы. Редактирование файлов осуществляется с учетом принятых для каждого файла стандартов спецификаций (форматы файлов см. “ОИОС01”).

2.1.3. Программная компонента «SETIES».

2.1.3.1. Структура и функциональные возможности программного компонента «SETIES»

Программный комплекс «SETIES» предназначен для исследования генных сетей, динамических эпигенов и клеточных ансамблей на основе метода обобщенных пороговых моделей.

Задача расчета динамики генных сетей заданной структуры решается методом обобщенных пороговых моделей (ОПМ) (Tchurayev, 1991). Главная идея метода заключается в разделении молекулярной системы кодирующих полимеров и метаболитов на собственно управляющую подсистему и управляемую подсистему, причем управляющая подсистема описывается на языке дискретной математики, а управляемая – на языке теории дифференциальных уравнений с особыми правыми частями.

В методе ОПМ элементам эукариотической управляющей генной сети сопоставляются формальные *генетические блоки*. Входными переменными генетического блока являются концентрации регуляторных веществ для соответствующего гена, а выходами – переменные, отражающие концентрации продуктов гена (РНК и белка). Информационную микроструктуру генетического блока образуют дискриминаторы D_{ij} , элементарные подавтоматы λ_{ij} ,

комбинационная схема S_j и элемент временной задержки TI_j . В дискриминаторе D_j величина концентрации r_j регуляторного белка сравнивается с пороговым значением P_j , при достижении которого происходит эффективное белок-ДНК взаимодействие. Элементарный подавтомат λ_j описывает процесс связывания регуляторного белка с операторным участком. Комбинационная схема S_j в зависимости от типа регуляторных отношений вырабатывает интегральную реакцию при воздействии разных регуляторных молекул на функционирование данного генетического блока. Элемент TI_j учитывает временную задержку, необходимую для транскрипции соответствующей нуклеотидной последовательности и созревания молекулы мРНК. Дискретный логический элемент L_j , состоящий из элементарных подавтоматов λ_j и комбинационной схемы S_j , в зависимости от концентраций регуляторных веществ $x_{i3}(t)$ формирует значения управляющей переменной $u_j = u_j(t)$ с учетом временной задержки (см. рис. 2.1.3.-1).

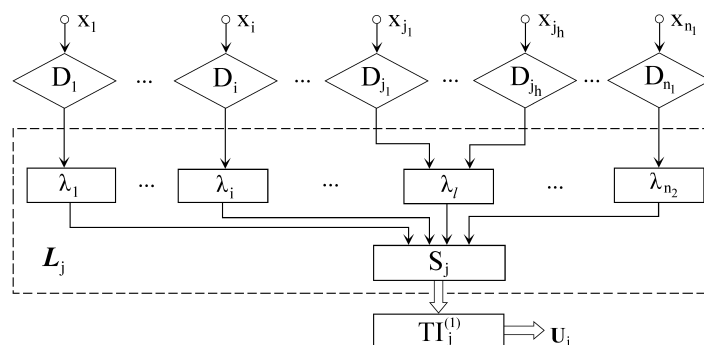


Рис. 48. Информационная микроструктура генетического блока. Пояснения в тексте.

Управляющая переменная определяет режим функционирования исполняющего устройства, которое описывает механизмы синтеза мРНК и белка и задается системой кусочно-линейных дифференциальных уравнений следующего вида:

$$\begin{aligned} \frac{\partial m_j^{(0)}(t)}{\partial t} &= a_{1j} u_j(t) - b_{1j} m_j^{(0)}(t), \quad a_{1j} = \sum_{l=1}^{k_j} a_{lj} u_{lj}(t), \\ \frac{\partial r_j(t)}{\partial t} &= a_{2j} m_j(t) - b_{2j} r_j(t), \end{aligned}$$

где a_{lj} – “удельная единичная сила” одного (l -го) из k_j сайтов специфичности данного блока; $u_{lj}(t)$ – управляющая дискретная переменная – компонента управляющего вектора; $m_j(t)$ – текущая концентрация, измеренная числом молекул на клетку, j -го транскрипта; a_{2j} – единичная интенсивность трансляции с j -го транскрипта; b_{1j} и b_{2j} - коэффициенты деградации j -го транскрипта и j -го полипептида соответственно; $r_j(t)$ – концентрация (в молекулах на клетку) синтезированного в блоке полипептида (выход системы).

Понимание текста размещенного в разделе 2.1.3. подразумевает, что пользователь ознакомился со стандартом и форматом файлов, применяемых в программном компоненте "SETIES". Подробное описание стандарта и формата данного файла дано в “ОИОС01”. Подробное описание руководства пользователя дано в “АСНИ-01”.

2.1.3.1. Создание компьютерных моделей молекулярно-генетических систем в программной компоненте SETIES.

Запуск SETIES

Для запуска программы кликните иконку «SETIES 5.0» на рабочем столе. При входе в систему на экране отображается рабочее окно в системе (см. рис. 2.1.3.-2).

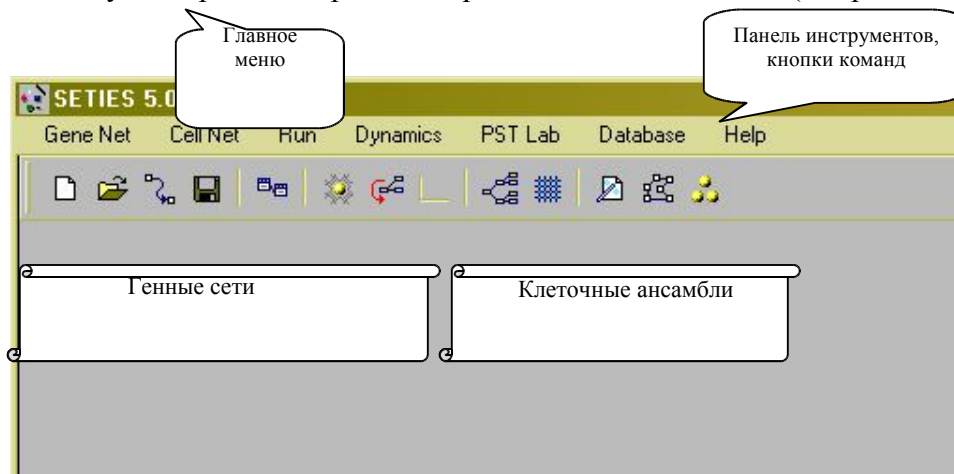


Рисунок 49. Рабочее окно системы SETIES.

Регистрация базы данных

Зарегистрируйте в программе «BDE Administrator» (установлена вместе с SETIES), выполнив следующую последовательность шагов:

Нажав кнопку «Пуск» меню «Windows», выберите «BDE Administrator» в списке установленных программ. В окне программы откроется список доступных баз данных (Рис. 50).

Активируйте пункт меню «Object → New». В появившемся окне укажите тип драйвера базы данных как «STANDART». В списке *псевдонимов* баз данных на левой панели появится соответствующий элемент (Рис. 50). Переименуйте поле «STANDART1» в «dbGeneNT».

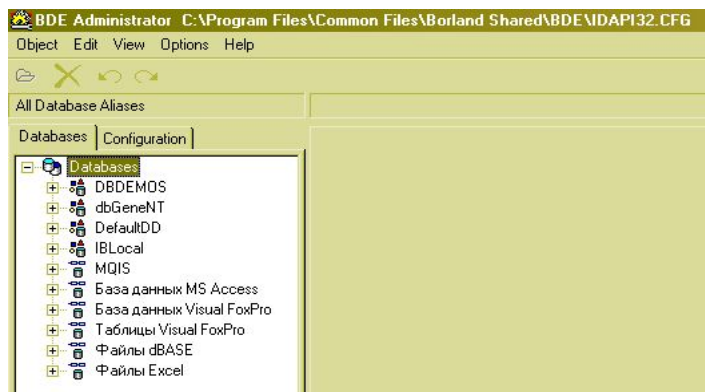


Рисунок 49. Список доступных баз данных.

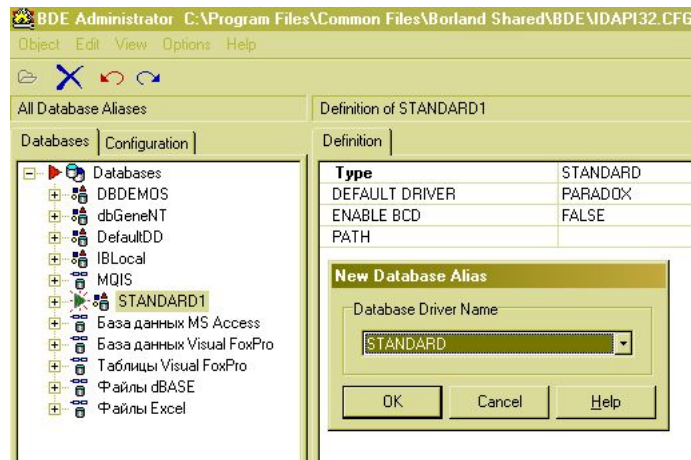


Рисунок 50. Поле STANDART1.

В строке «PATH» в таблице определения укажите путь к каталогу с файлами базы данных: «C:\Program Files\MMG\SETIES 5.0\dbGeneNT» (эта папка была создана ранее программой установки SETIES).

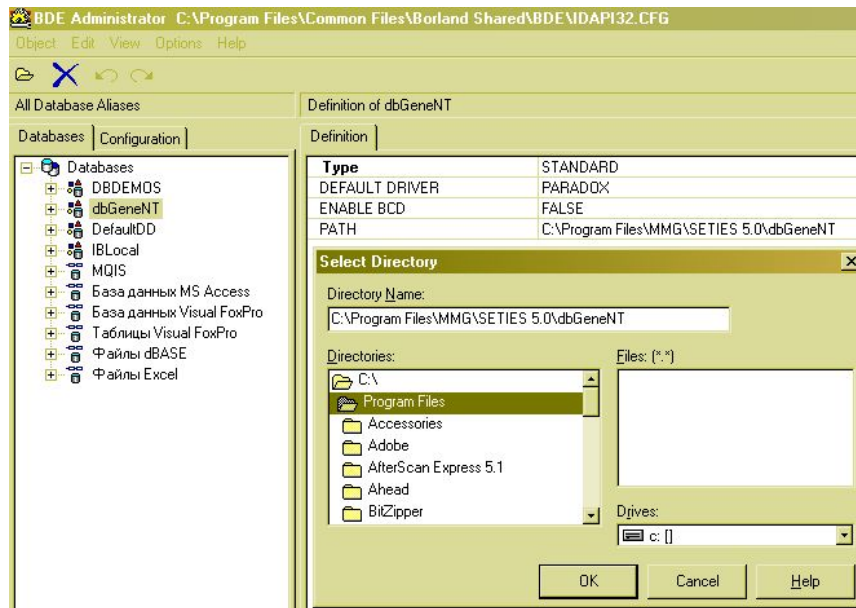


Рисунок 51. Путь к каталогу с файлами базы данных.

Сохраните изменения, выбрав пункт меню «Object → Apply», и выйдите из «BDE Administrator».

Создание информационной модели генной сети

Возможны четыре способа привнесения в среду программы информационная модель генной сети (ИМГС), по которой далее автоматически строится математическая модель на языке ОПМ-формализма:

- 1) Загрузите встроенные в программу демо-примеры.
- 2) Выберите запись из базы данных. База данных размещается в папке «C:\Program Files\MMG\SETIES 5.0\dbGeneNT».

- 3) Соберите по make-файлу (текстовый файл специального формата). Входящие в инсталляционный пакет пять файлов сборки размещаются в папке «C:\Program Files\MMG\SETIES 5.0\Makefiles».
- 4) Сгенерируйте посредством «Конструктора генных сетей».

Загрузка демо-примеров

Выполните *Главное меню: Help → Samples → xxx*. При этом программа загрузит одну из семи генных сетей, реализованных в исходном коде (Рис. 52), в точности:

- циклическую моногенную систему (CMS);
- циклическую дигенную систему или молекулярный триггер (CDS);
- генную сеть, содержащую три CDS, или мультистат (TCDS);
- генную сеть на базе TCDS с добавлением одной связи (TCDS+);
- 6-компонентную генную сеть (Sixtomic);
- генную сеть на базе Sixtomic с добавлением одной связи, вызывающей режим осцилляции (Sixtomic Osc);
- генную сеть размером в триста G-блоков (Thousannah).

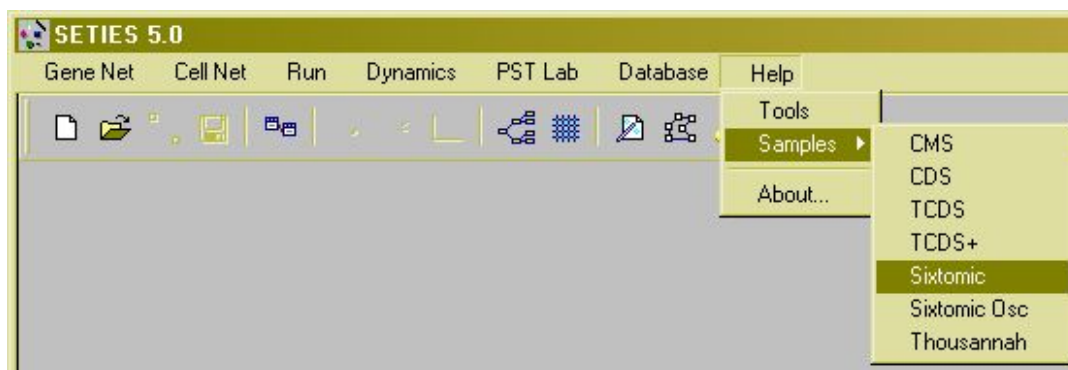


Рисунок 52. Загрузка демо-примеров.

Выборка записи из базы данных

Выполните *Главное меню: Gene Net → Load → Database* (Рис. 53).

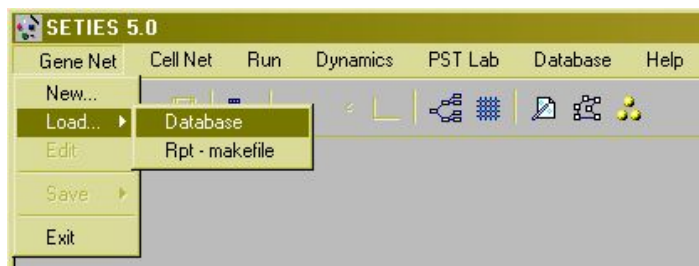


Рисунок 53. Выборка записи из базы данных.

Появляется окно со списком наименований генных сетей (ГС), хранящихся в базе данных (Рис. 54). Отметьте курсором строку с наименованием нужной системы и нажмите кнопку «Load». При этом произойдет автоматическая загрузка модели из таблиц базы данных в программный объект – информационную модель генной сети.

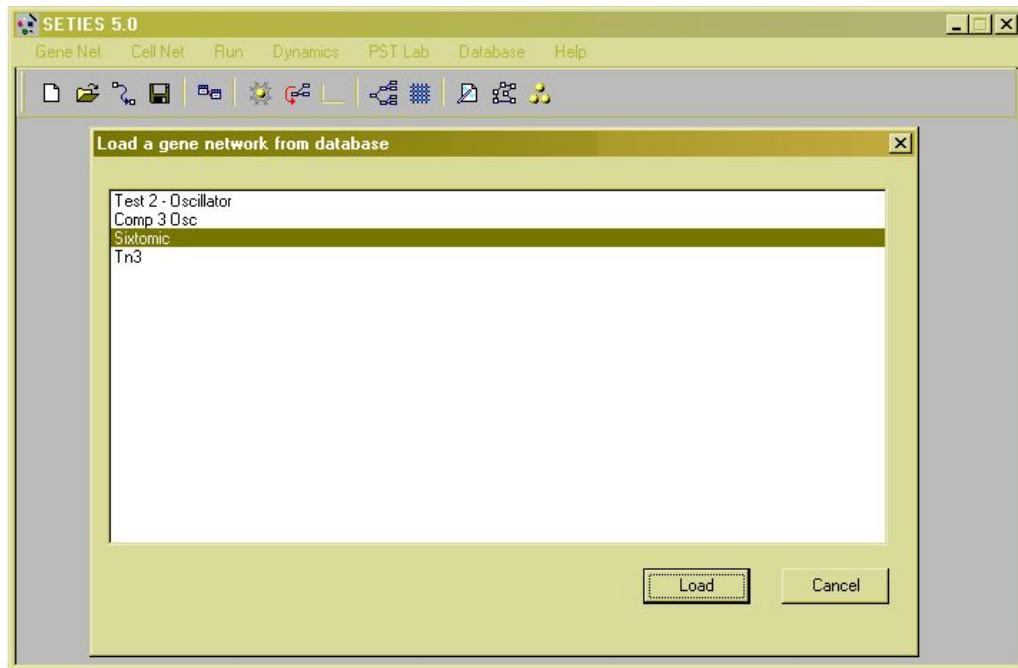


Рисунок 54. Выборка наименований нужной системы.

Сборка информационной модели по make-файлу

Выполните *Главное меню: Gene Net → Rpt – makefile.*

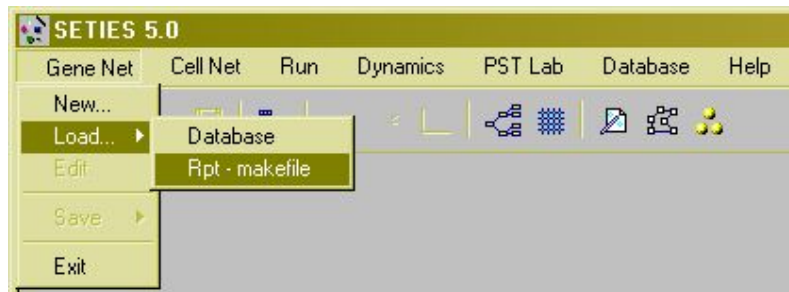


Рисунок 55. Сборка информационной модели по make-файлу.

Программа откроет диалоговое окно с предложением указать файл сборки (Рис. 56).

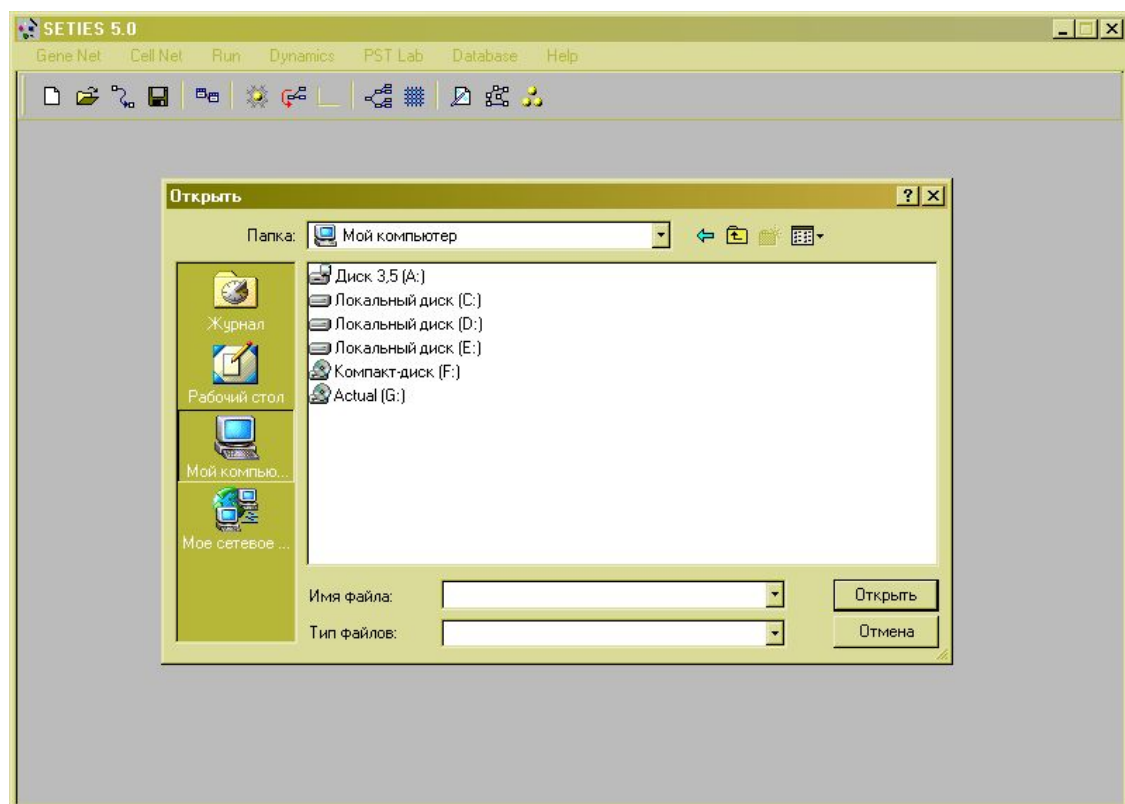


Рисунок 56. Диалоговое окно для поиска файла сборки.

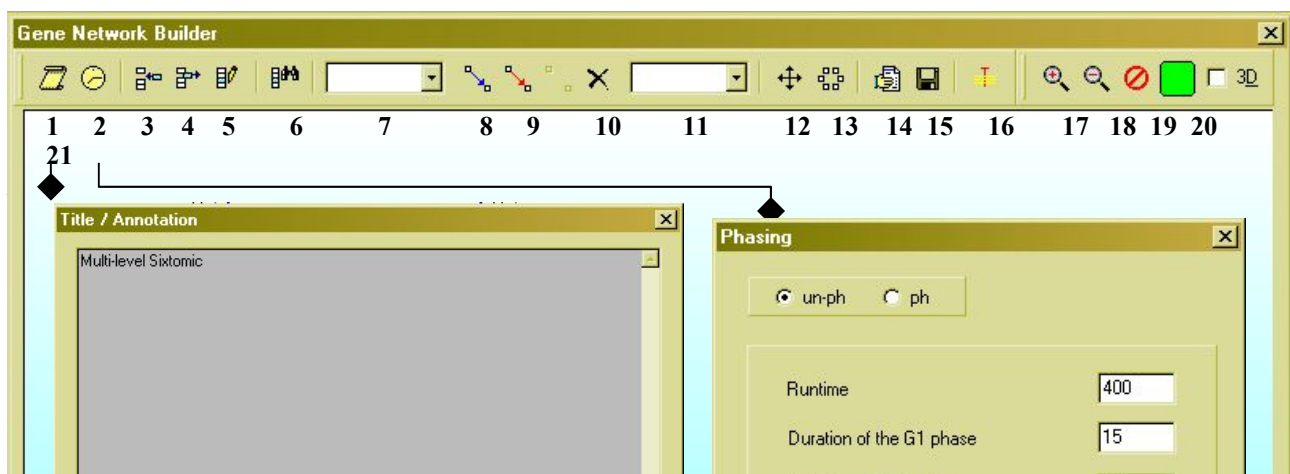
Вызов «Конструктора генных сетей»

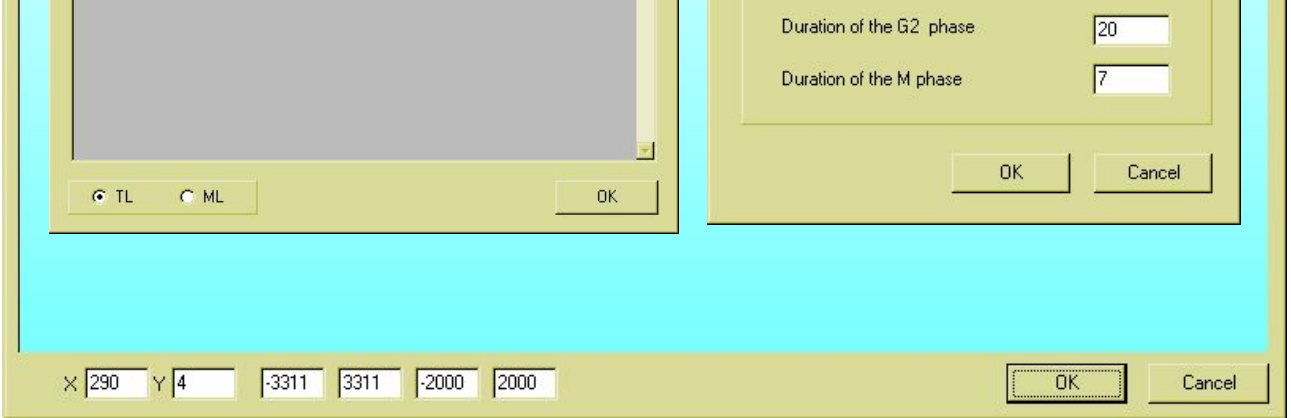
Данная процедура выполняется в двух режимах:

- в режиме первичной сборки (*Главное меню: Gene Net → New...*), при этом данные о предыдущем объекте, если тот существовал, теряются;
- в режиме редактирования (*Главное меню: Gene Net → Edit...*), который возможен при наличии в среде программы экземпляра ранее загруженной системы (например, одним из трех вышеописанных способов).

Все операции этого процесса выполняется посредством «Конструктора генных сетей» (Рис. 57). Для каждого объекта, будь то генная сеть, генетический блок (g-блок), регуляторная связь (r-связь), регуляторный модуль (r-модуль), регуляторный сайт (r-сайт), набор констант синтеза - поддерживается сервис "откат к резервной копии". Таким образом, при необходимости можно отменить сделанные изменения и вернуться к исходному набору значений данных.

Системный уровень





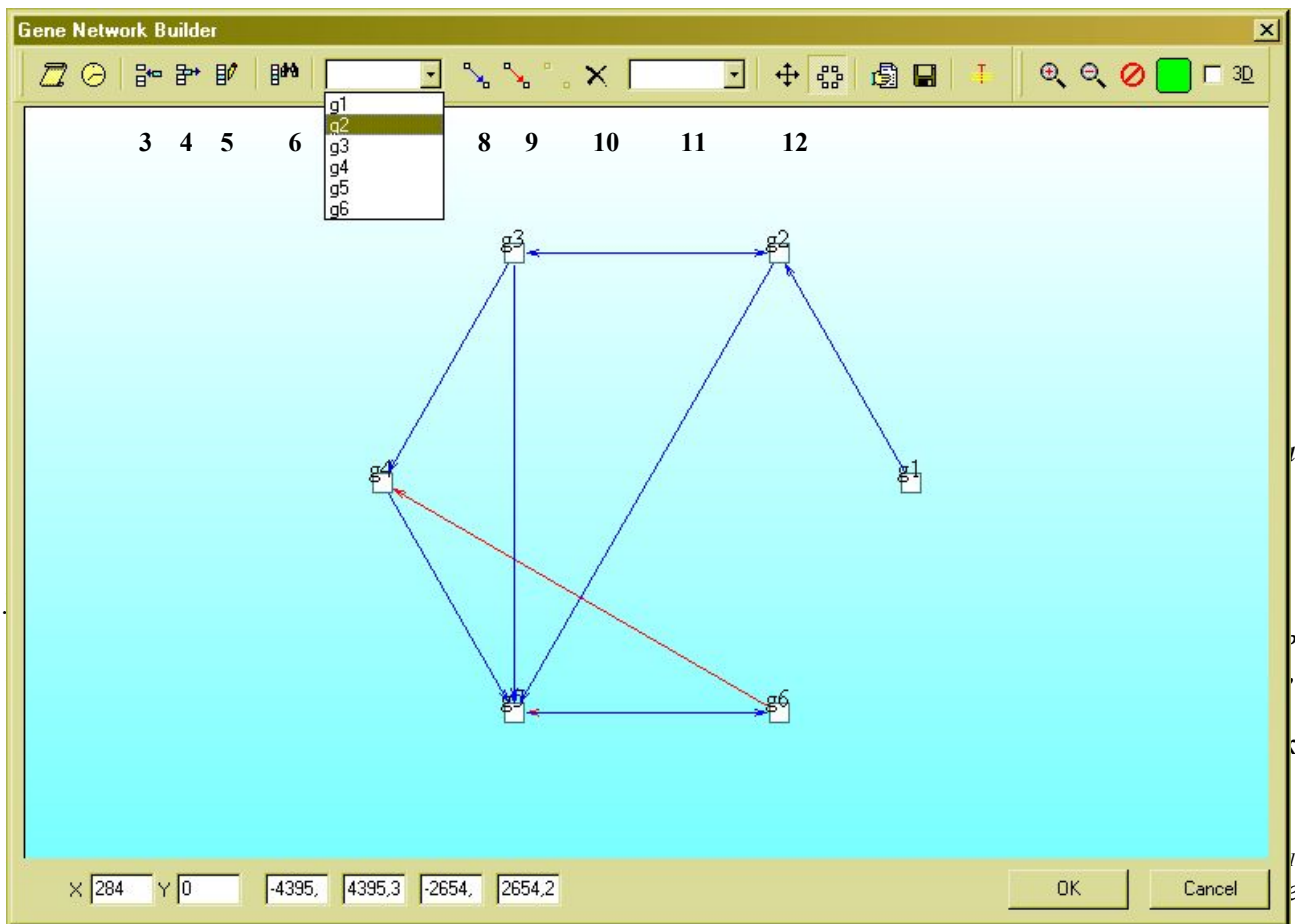
12, 13 – выбрать режим размещения элементов ГС на граф-схеме: произвольный (12), по окружности (13, по умолчанию);

14 – генерировать отчет по входным данным: кликните правой кнопкой мыши – появится всплывающее меню с двумя пунктами (документ или исполнимый модуль);

15 – сохранить граф-схему ГС в файле *.bmp;

16 – загрузить ГС из БД;

17-21 – работать с изображением: увеличить (17), уменьшить (18), восстановить (19), поменять фон (20), переключиться между 2D и 3D форматом(21).



завершено, то другие действия с другими элементами блокируются. нажмите кнопку 8, 9 или 10 и кликните на втором элементе.

Те же процедуры можно выполнять в пределах панели (например, для больших ГС), используя выпадающие списки. В этом случае изображению элемента на схеме соответствует пункт списка с наименованием g-блока. Кнопки 3-6 ориентированы на список 7. Например, чтобы найти g-блок на холсте:

нажмите кнопку 6 и сделайте выбор в списке 7 (искомый элемент подсветится).

Чтобы добавить (удалить) связь:

выберете g-блок из списка 7,
нажмите кнопку 8 или 9 (10) и
выберете g-блок из списка 11.

Генетические блоки: кинетические параметры

Сделайте выбор поля, соответствующее какому-либо параметру, и задайте с клавиатуры допустимое значение.

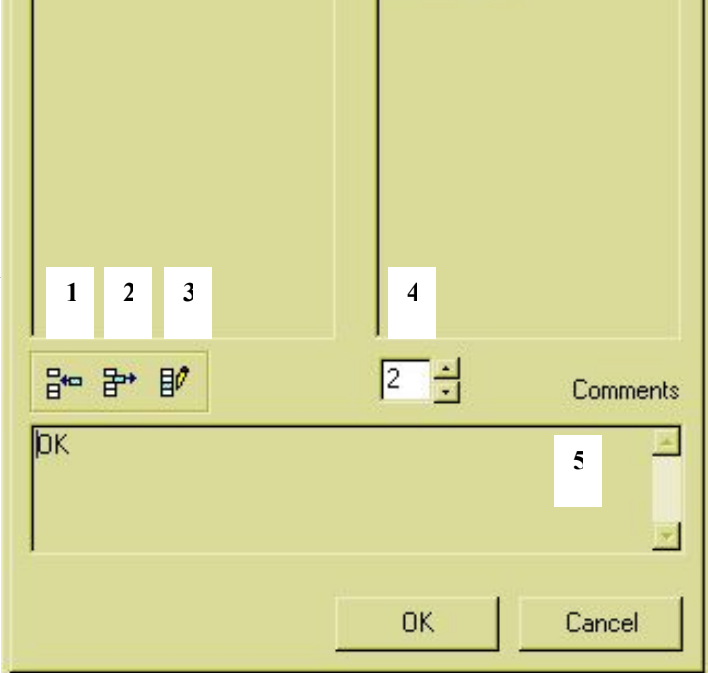
The screenshot shows two main windows. The left window is titled 'G-block Data-in' and has three tabs: 'Parameters', 'Interactions', and 'Modules'. The 'Parameters' tab is active, showing fields for 'gbID' (5), 'gbName' (g5), 'Proper State' (nonactive), and 'Site Count' (5). The 'Interactions' tab is also visible, showing a table with columns 'rbID', 'Upstream Gene', 'rbType', 'Threshold', and 'R-period'. The table contains five rows of interaction data. Below the table are buttons '1' and '2', and a 'Logic Function, (P)DNF' field with the expression 'X1x3+X1X2x3'. The right window is titled 'The Synthesis' Rates' and has a 'Select Genes' dialog open. The dialog shows a list of genes: '1 g1', '2 g2', '3 g3', '4 g4', '5 g5', and '6 g6'. Gene '5 g5' is selected. There are 'OK' and 'Cancel' buttons at the bottom of the dialog.

копийность.

Для редакции доступны только g-модули из списка в левой панели. После редакции прототипов, все имеющиеся экземпляры g-модулей удаляются и вместо них нужно сформировать новые (просто задав копииность).

Выберете курсором строку с его названием и нажмите кнопку «2».

The screenshot shows the 'G-block Data-in' window with the 'Modules' tab selected. The 'Modules' tab shows a list of modules: 'module1', 'module2', 'module1 copy 1', 'module2 copy 1', 'module1 copy 2', and 'module2 copy 2'.

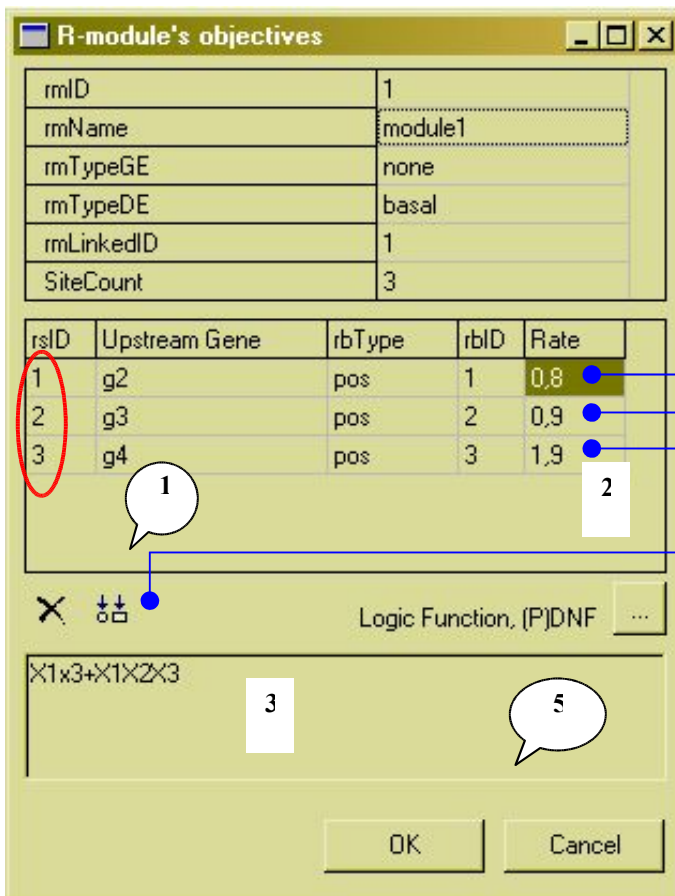


ков: регуляторных модулей.

«Select Interactions» кликните на строке, которая параметров связи в этом окне не из таблицы в окне «G-Block Data-in», лист

я окно «The synthesis' Rates», содержащем азах клеточного цикла.

логической функции - дизъюнктивной зоваться автоматизированным средством того осуществляется при нажатии кнопки



блоков: р

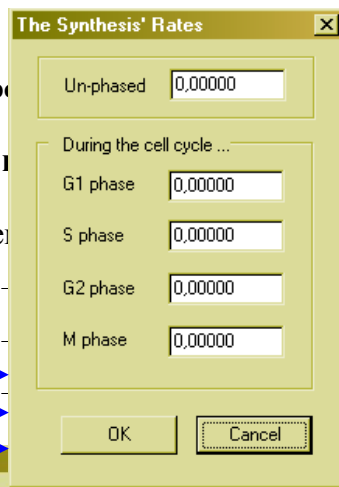
ой сети и

нить тре

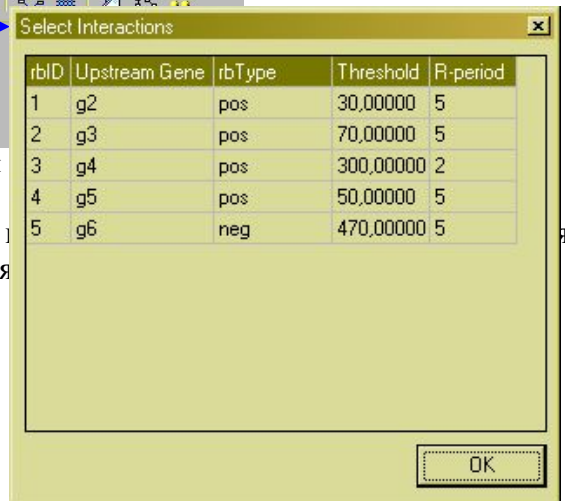
$e \rightarrow Rpt$

$e \rightarrow Rpt$

$e \rightarrow Rpt$



ИТОВ



ЯХ

ЛИ

Я

ДЛЯ

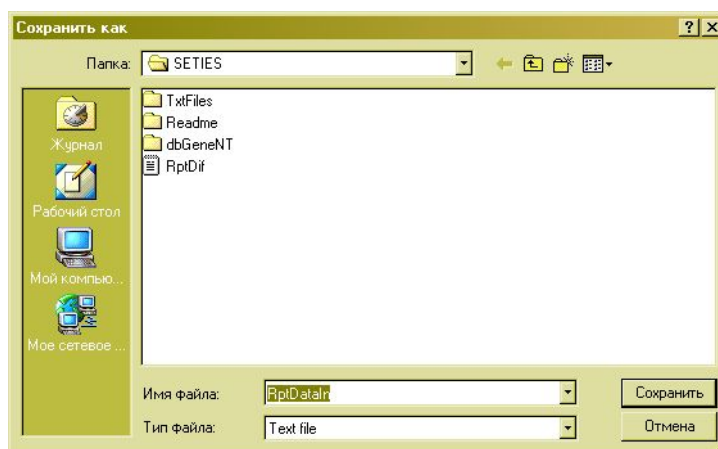


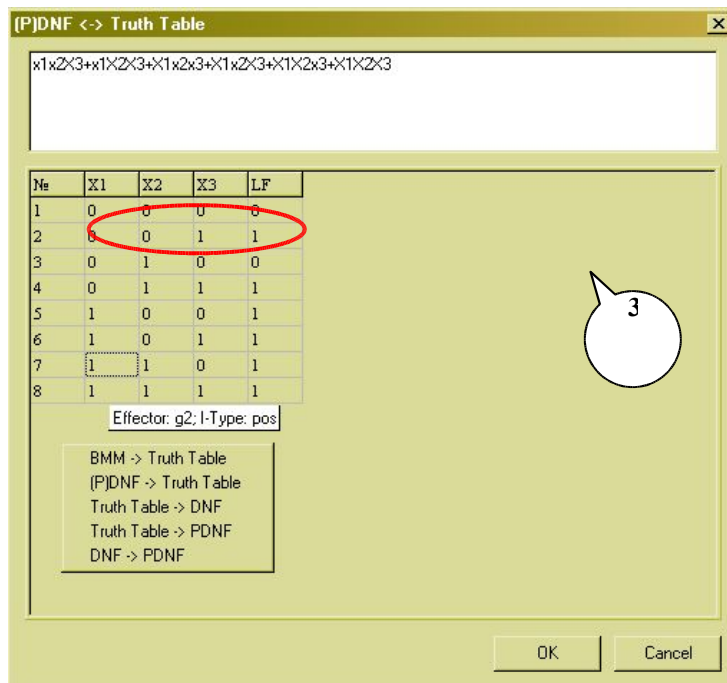
Рис. 64. Сохранение информационной модели генной сети в журнале экспериментов.

Корректирование входов задачи

Нажмите ОК в окне ГС-конструктора (Рис. 57).

Если значения параметров не соответствуют своим областям определения или есть пустые поля и т.д., то появляется сообщения ошибок ввода и приостанавливается переход на следующий уровень. В каждом сообщении содержится адрес ошибки (системный уровень, g-блок, g-связь, g-модуль, g-сайт), тип ошибки и рекомендация по ее устранению.

За согласованность (непротиворечивость) входных данных отвечает ДНФ-отладчик используется для автоматизации описания логики регуляторных механизмов на уровне g-блоков и g-модулей.



ДНФ-отладчика.

построение таблицы истинности (ТИ) по нормальной форме (ДНФ), построение ДНФ из совершенной (С) ДНФ (минимизация ДНФ). Для заполнения ТИ используется базовая логика, что: если концентрация хотя бы одного регулятора; если концентрация хотя бы одного регулятора.

заданной в текстовом редакторе

поле текстового редактора «3».

le.

ую далее можно редактировать.

При вводе ДНФ в редакторе «3» следует учитывать, что X_i – логическая переменная, соответствует i -му регуляторному сайту (на рис. 62 эти участки выделены овалом); № -



номер набора значений \bar{X} ; LF – значение функции (1 - вклад комбинации регуляторных сигналов в активацию транскрипции ненулевой, 0 - нулевой).

Построение ДНФ по значениям функции в ТИ

Заполните ТИ для искомой функции (вручную или с помощью шаблона быстрого ввода).

Кликните правой кнопкой мыши.

Выполните *Всплывающее меню: Truth Table → DNF*.

Полученный результат отобразится в текстовом редактора «3».

Расчет динамики в математических моделях

Запуск расчетов

Выполните *Главное меню: Run → Gene Net (eTL / eML + un-ph)*. Программа рассчитает динамику системы без учета фаз клеточного цикла и делений клеток.

Выполните *Главное меню: Run → Gene Net (eTL / eML + ph)*. Программа рассчитает динамику системы с учетом фаз клеточного цикла и делений клеток.

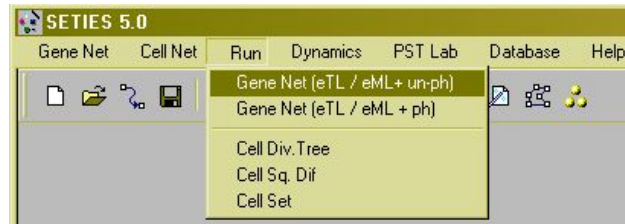


Рисунок 66. Запуск расчетов.

Просмотр результатов

Выполните *Главное меню: Dynamics → Gene Net*.

Появляется окно выбора профиля просмотра выходных данных (Рис. 67).

Выберете в списке «Genes» подмножество генов, для которых вы планируете просмотреть динамику активности, кликнув мышью на log-поле при наименовании того или иного гена.

Выберете в списке «Type of data-out» тип выходных данных (кинетические кривые для РНК, белков или диаграммы активности генов) и стиль демонстрации (автономный, сопоставление), кликнув мышью в соответствующих разделе и строке.

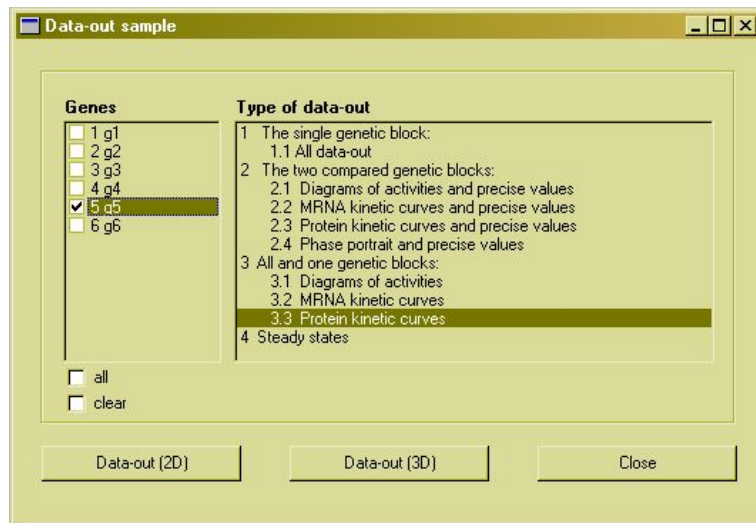


Рисунок 67. Выбор профиля просмотра выходных данных.

Нажмите кнопку «Data-out (2D)» или «Data-out (3D)».

Появляется окно с выходами системы в соответствии с запрошенным профилем (Рис. 68).

На графике по оси Ox – время (в тактах), по оси Oy – значения измеряемой величины, (концентрации продуктов генов, выраженные в молекулах на клетку, уровни активности генов), по Oz – наименования генов.

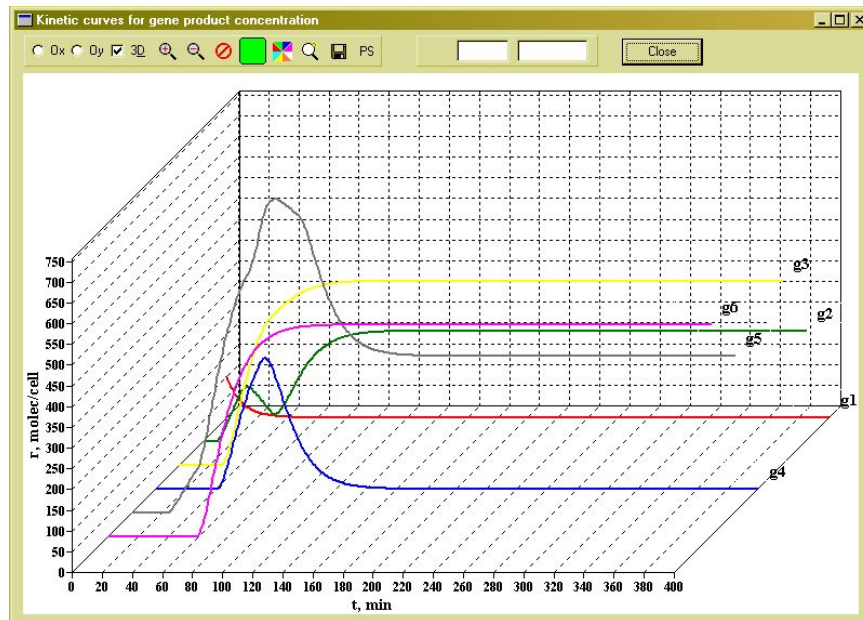


Рисунок 68. Результаты расчёта.

2.1.4. Программная компонента «STEP+»

2.1.4.1. Структура и функциональные возможности программного компонента «STEP+»

Программный компонент «STEP+» ориентирован на комплексное исследование автономных систем уравнений, их стационарных решений и устойчивости стационарных решений в зависимости от параметров, а также нелинейных систем общего вида. В таком виде формулируются многие математические модели химической кинетики, физической химии, математической биологии и других приложений.

Понимание текста размещенного в разделе 2.1.4. подразумевает, что пользователь ознакомился со стандартом и форматом файлов, применяемых в программном компоненте «STEP+». Подробное описание стандарта и формата данного файла дано в “ОИОС01”. Подробное описание руководства пользователя дано в “АСНИ-01”.

2.1.4.1.1. Создание компьютерных моделей в программном компоненте «STEP+»

Запуск STEP+

Запустите исполняемый файл create.exe в директории, в которой размещен STEP+. Появляется диалоговое окно (Форма Start) (Рис. 69):

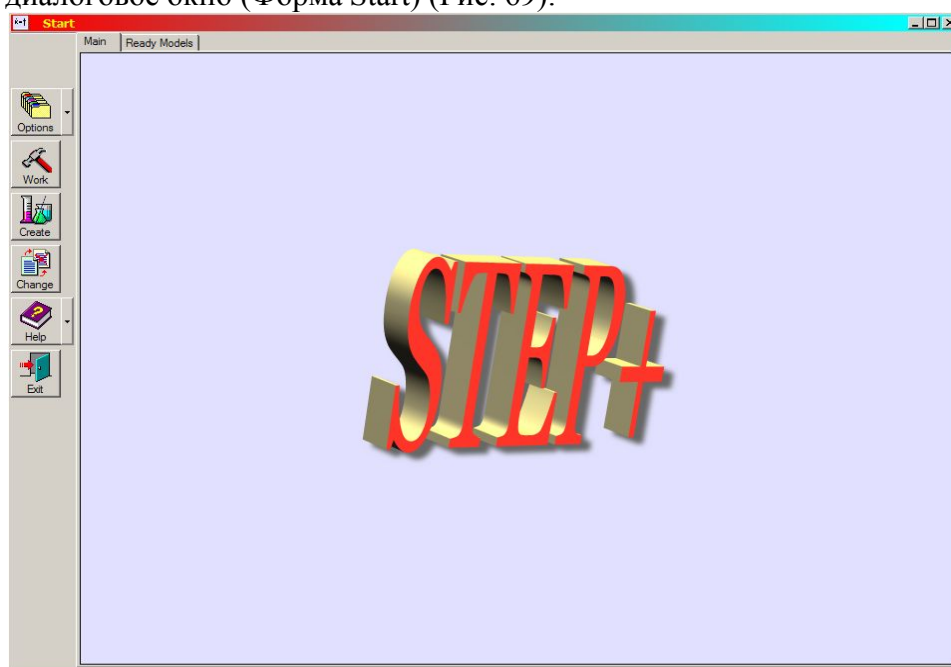




Рисунок 69. Форма Start.

Создание новой модели

Нажмите на одну из двух кнопок на стартовой форме: кнопку 'Create'  или кнопку 'Change' . После загрузки формы 'Creation' во второй субпанели её панели состояния выводится название субрежима создания ('Create' или 'Change'):

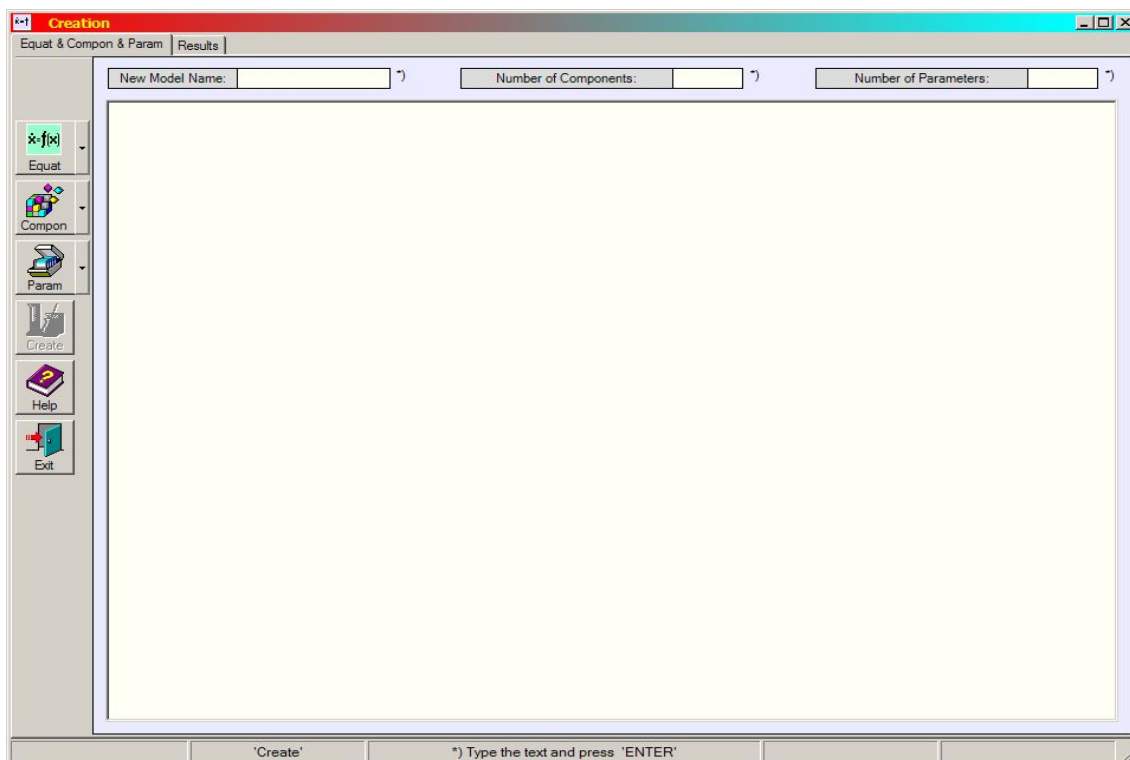


Рисунок 70. Создание новой модели.

Введите имя новой модели. Нажмите кнопку 'ENTER' на клавиатуре

Введите правые части системы дифференциальных уравнений. Нажмите кнопку 'ENTER' на клавиатуре.

Введите число компонент (число уравнений в системе). Нажмите кнопку 'ENTER' на клавиатуре.

Введите число параметров модели. Нажмите кнопку 'ENTER' на клавиатуре.

Работа с готовой моделью

Выберите модель из списка моделей (стартовая форма, вкладка 'Ready Models').

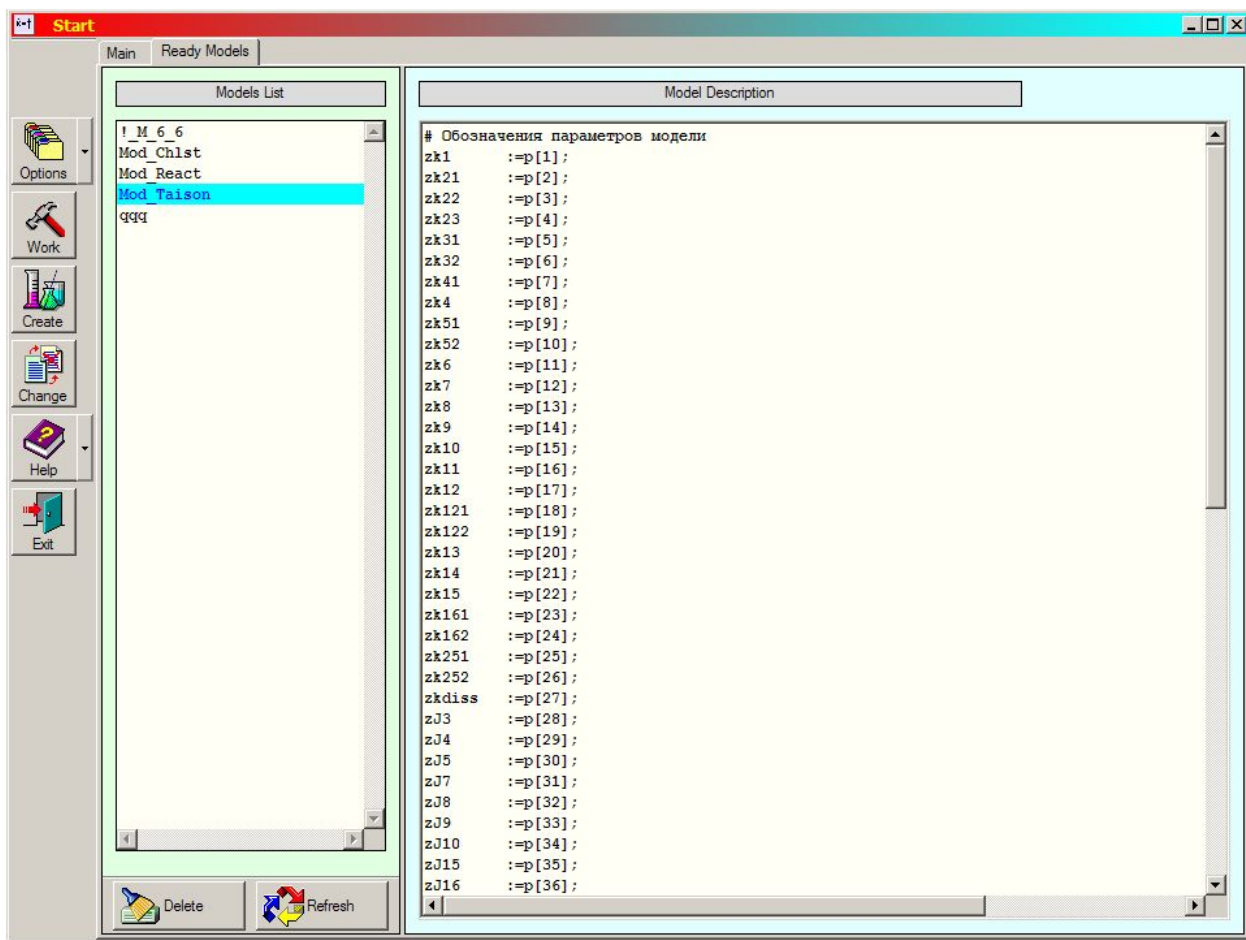



Рисунок 71. Список готовых моделей.

Нажмите кнопку 'Work' . Загружается форма 'Working', на которой в первой субпанели панели состояния отображается имя выбранной модели и открыта первая вкладка второго ряда 'Components & Parameters', принадлежащая первой вкладке первого ряда 'Integration':

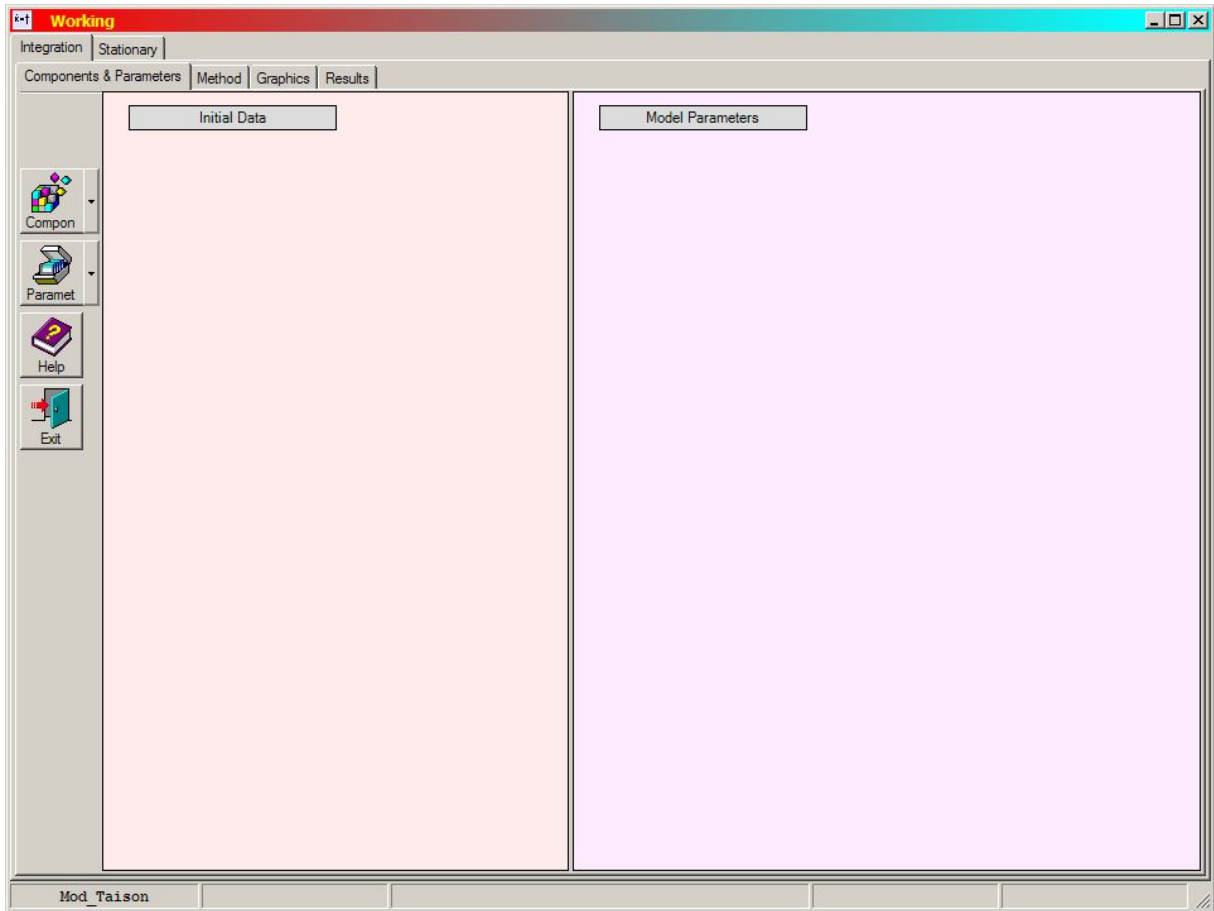



Рисунок 72. Вкладки модели.

Решение задачи Коши

Задайте начальные данные для компонент: нажмите кнопку 'Compon'  формы 'Working'. Появляется диалоговое окно.

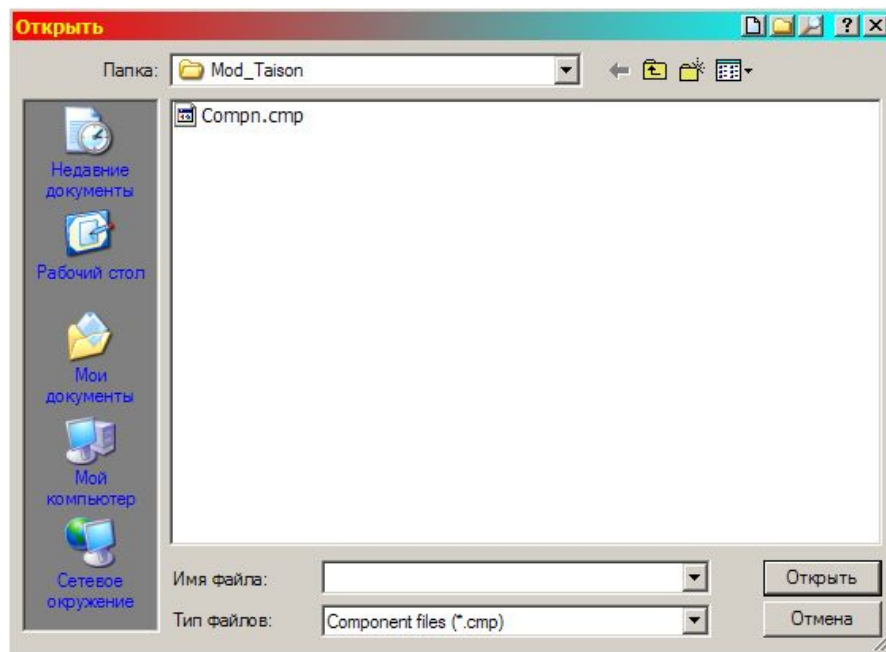



Рисунок 73. Диалоговое окно поиска файлов.

Выберите нужный файл и нажмите кнопку 'Открыть'. В панели 'Initial Data' появляется таблица с начальными данными:

Initial Data	
$x_0(1) =$	1.000000E-01
$x_0(2) =$	2.500000E-02
$x_0(3) =$	1.000000E+00
$x_0(4) =$	0.000000E+00
$x_0(5) =$	1.180000E+00
$x_0(6) =$	0.000000E+00
$x_0(7) =$	1.000000E-01
$x_0(8) =$	0.000000E+00

Задайте значения параметров модели: нажмите кнопку 'Paramet' . Появляется диалоговое окно:

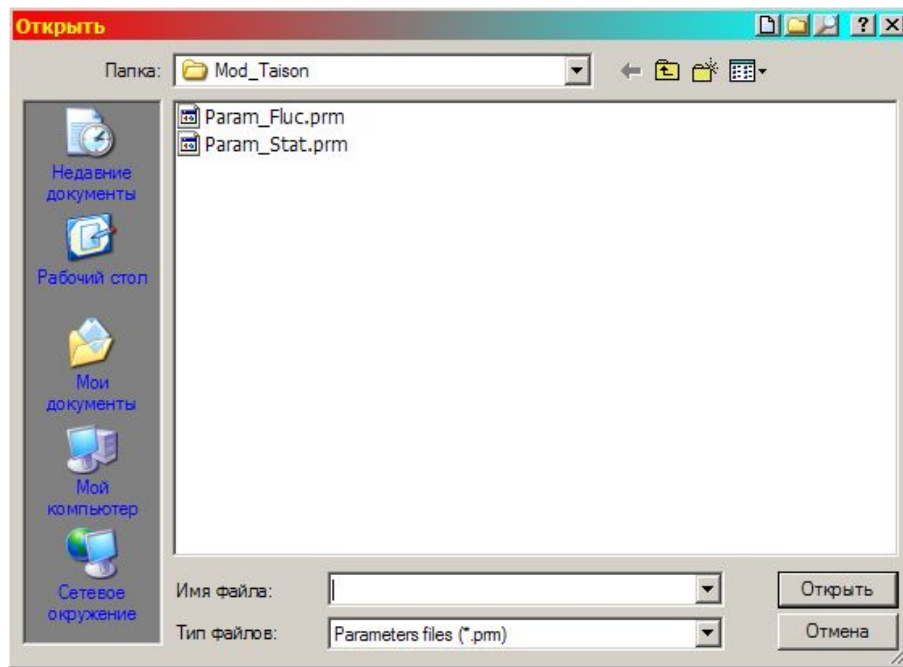


Рисунок 74. Диалоговое окно поиска файлов.

Задайте значения параметров метода: перейдите на вкладку 'Method', нажав левой кнопки мыши на имя вкладки:

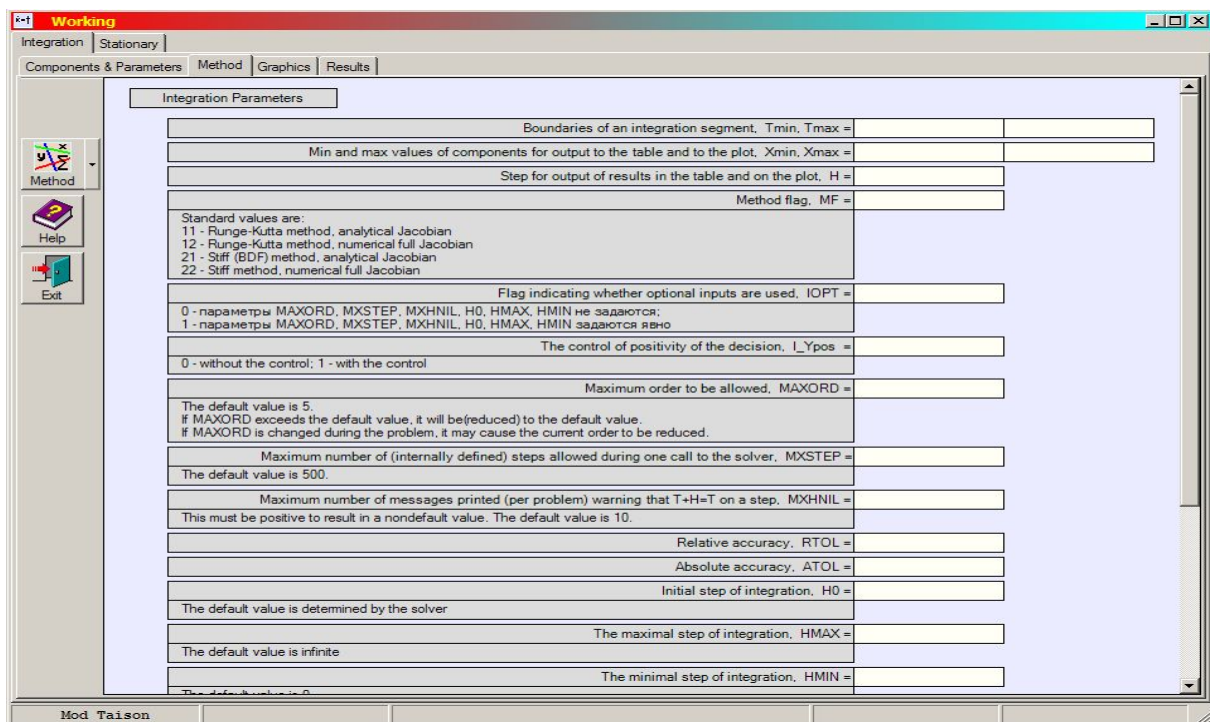



Рисунок 74. Вкладка Method.

После этого необходимо нажать левой кнопки мыши на кнопку 'Method'  - при этом появится следующее диалоговое окно:

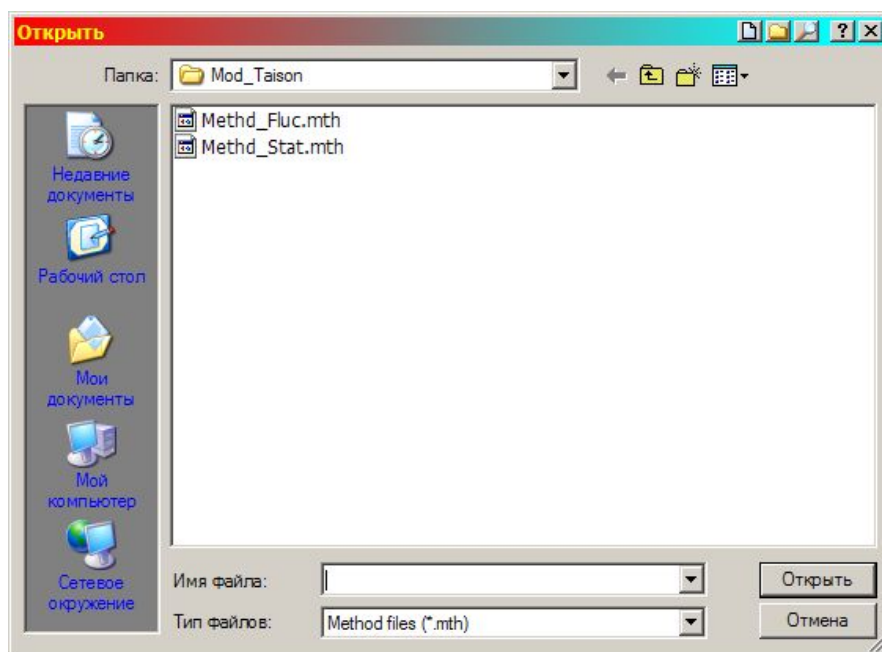


Рисунок 75. Диалоговое окно поиска файлов.

Выберите файл 'Method_Fluc.mth' двойным щелчком на имени метода.

В результате в рабочем пространстве пакета появятся текстовые поля панели 'Integration Parameters', которые заполняются числовыми значениями:

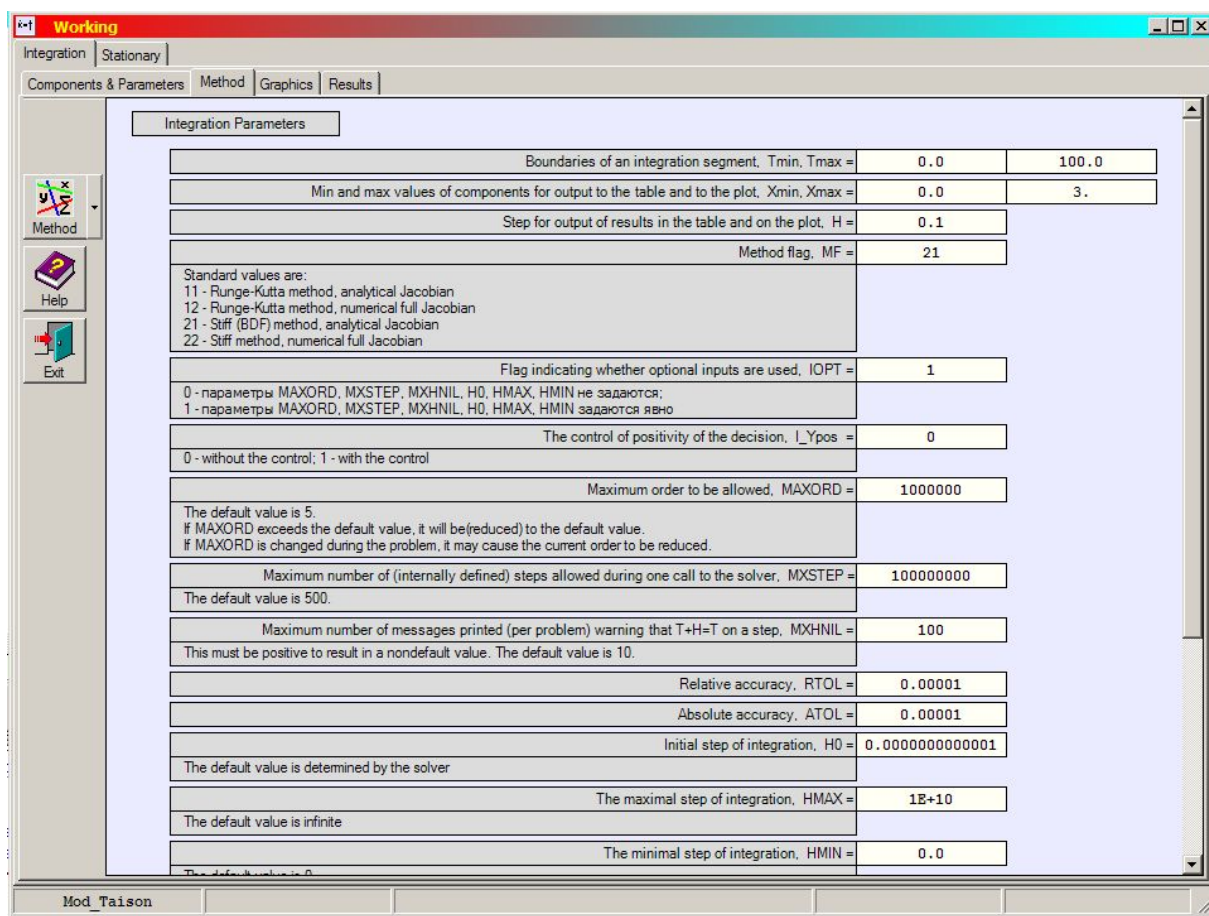


Рисунок 76. Панель Integration Parameters.

Запуск на расчёт. Визуализация расчётов модели: перейти на вкладку 'Graphics', нажав на имя вкладки:

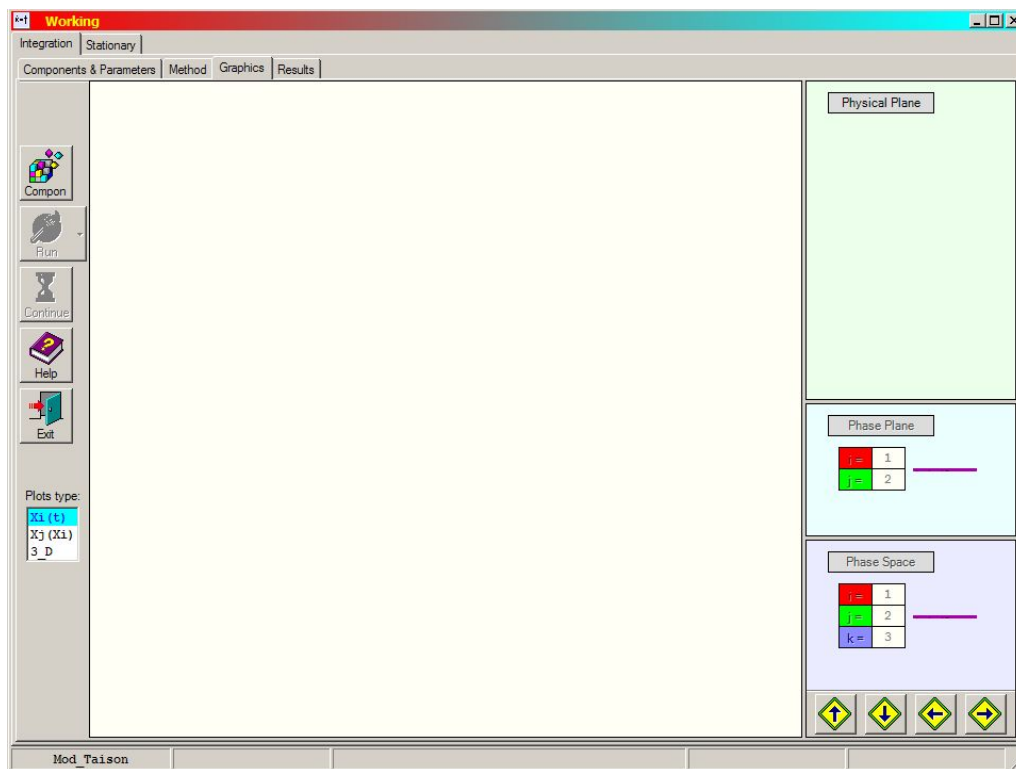



Рисунок 77 Вкладка 'Graphics'.

Нажмите на кнопку 'Compon' . В панели 'Physical Plane' появляются номера и цвета компонент, выводимых на график, и кнопка 'Run' становится активной:

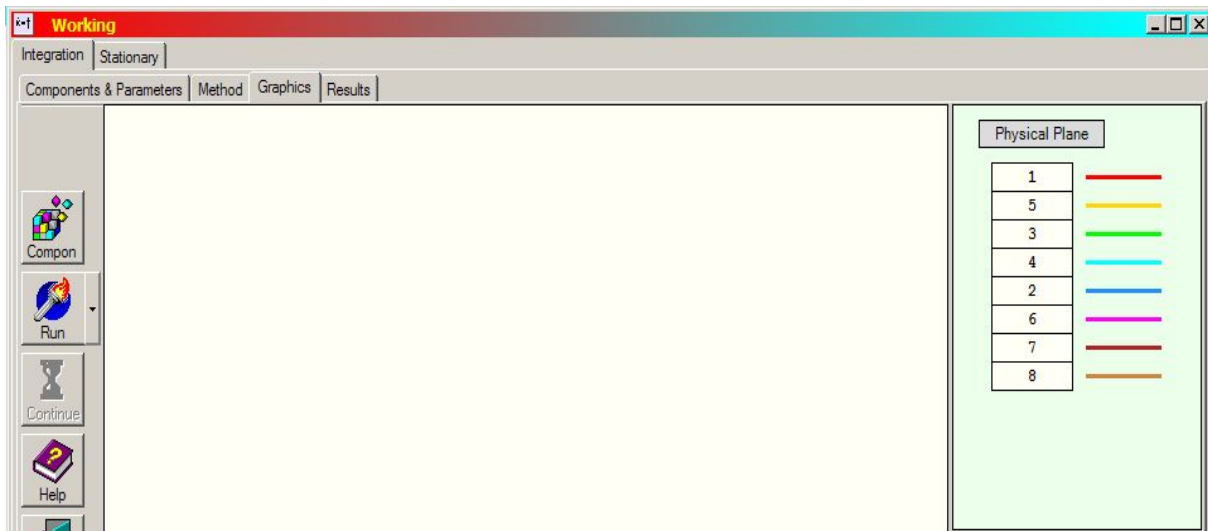


Рисунок 78. Появление активной кнопки 'Run'.

Нажмите на кнопку 'Run' .

Компоненты решения визуализируются в главном окне вкладки 'Graphics':

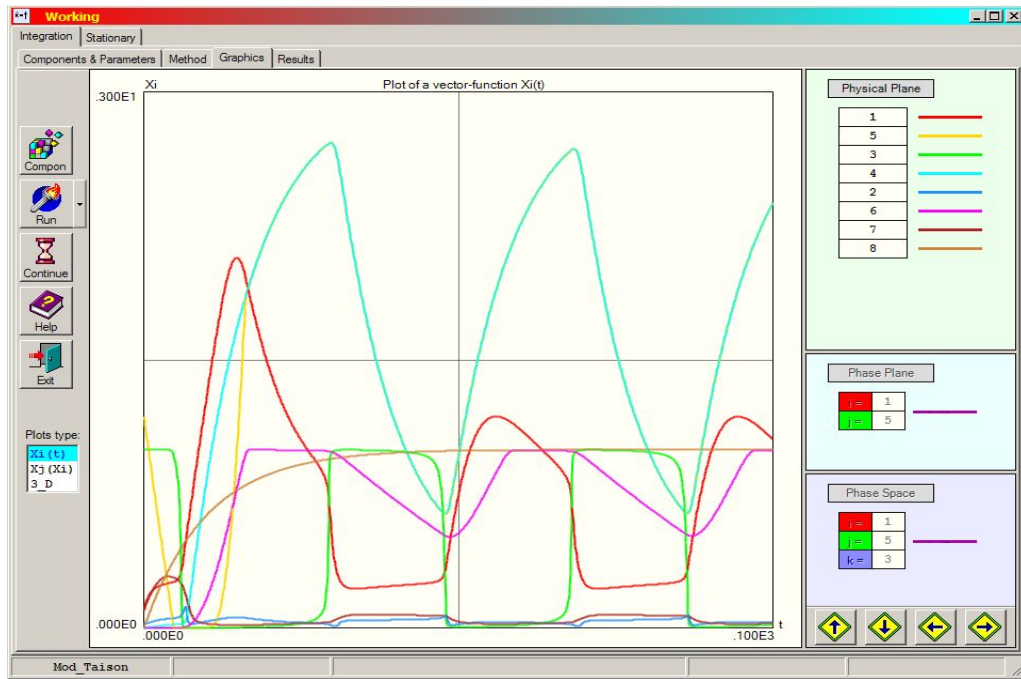


Рисунок 79. Визуализация расчётов модели.

Выберите в списке типов графиков 'Plots type' значение 'Xj(Xi)' и нажмите на кнопку 'Run'. В результате произойдёт расчёт модели с выводом графика для выбранных компонент в фазовой плоскости:

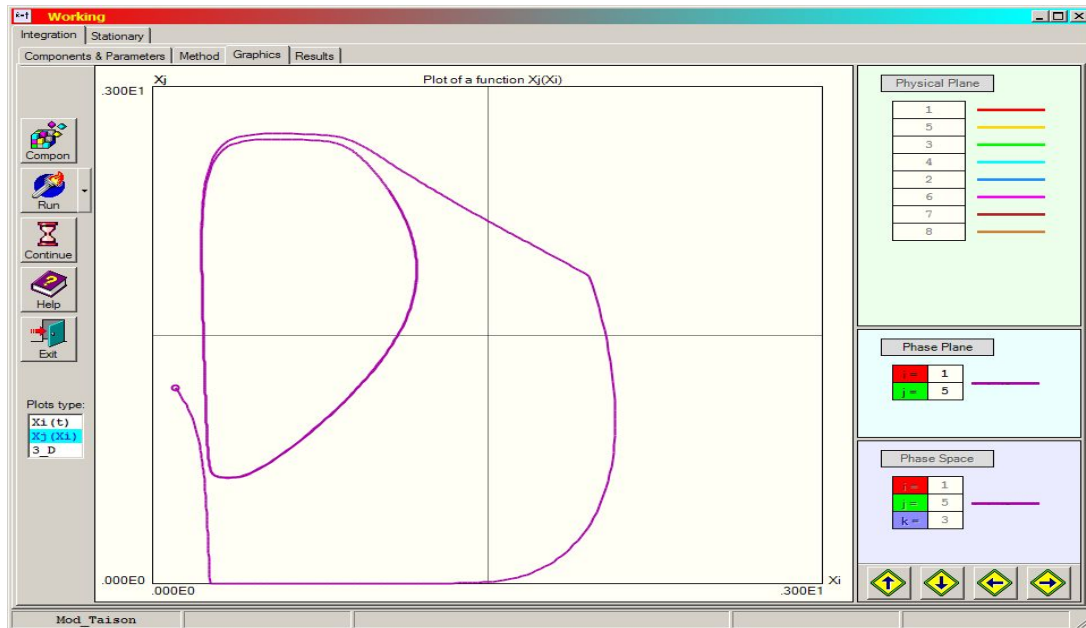


Рисунок 80. Визуализация расчётов модели в фазовой плоскости.

Выберите в 'Plots type' значения '3_D' и нажмите на кнопку 'Run'. В результате будет визуализирована пространственная картинка для выбранных компонент:

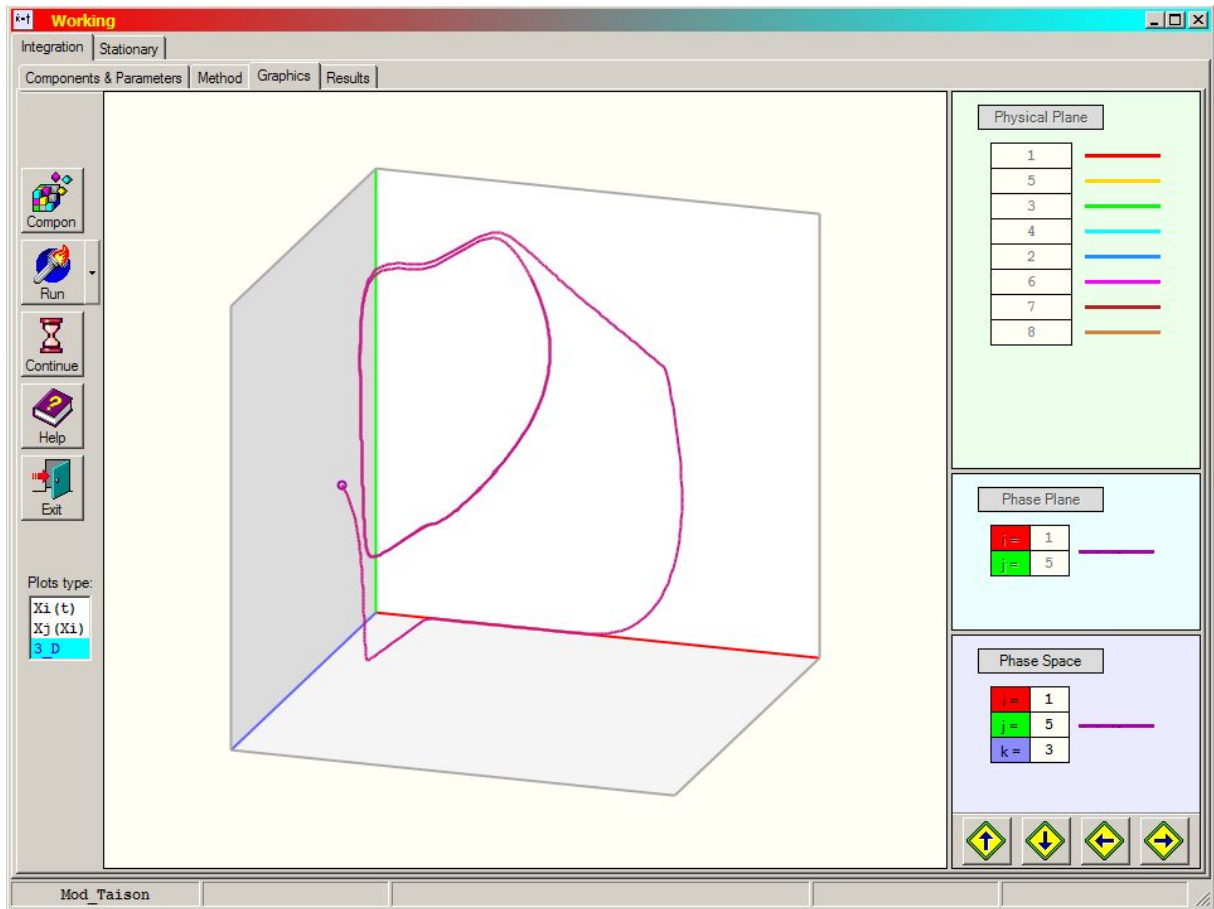



Рисунок 81. Визуализация результатов расчёта в трёхмерном пространстве.

Нажмите на кнопку 'Continue'  - решение (задачи Коши) продолжается на новый временной интервал, значение которого (T_{max}) можно изменять на вкладке 'Method' (начало нового интервала по-прежнему остаётся равным 0). При этом можно изменять и значения параметров.

Анализ результатов расчёта. Табличные данные: перейдите на вкладку 'Results':

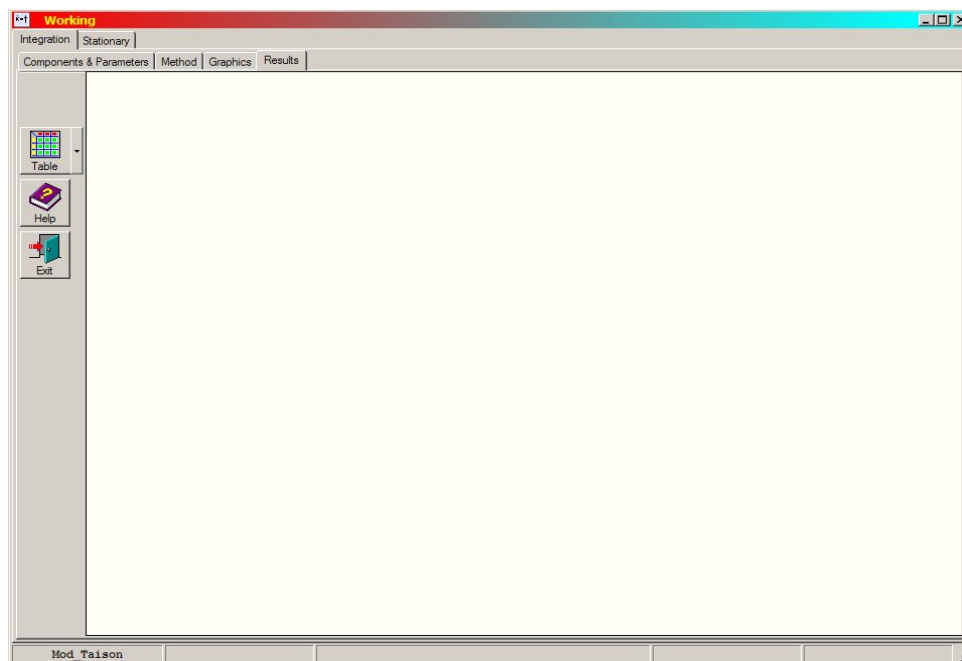



Рисунок 82. Вкладка 'Results'.

Нажмите на кнопку 'Table' . Появляется диалоговое окно:

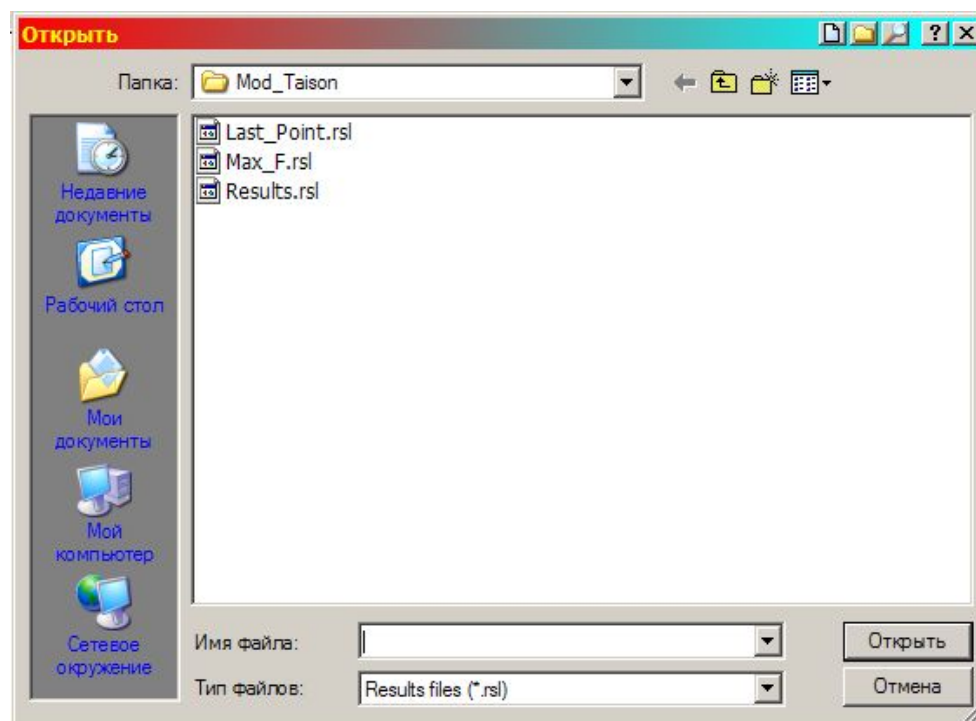


Рисунок 83. Диалоговое окно выбора файла различных типов результатов.

Исследование стационарного состояния

Получите стационарное решение в задаче Коши и выполните следующие действия:
Откройте вкладку 'Stationary' формы 'Working':

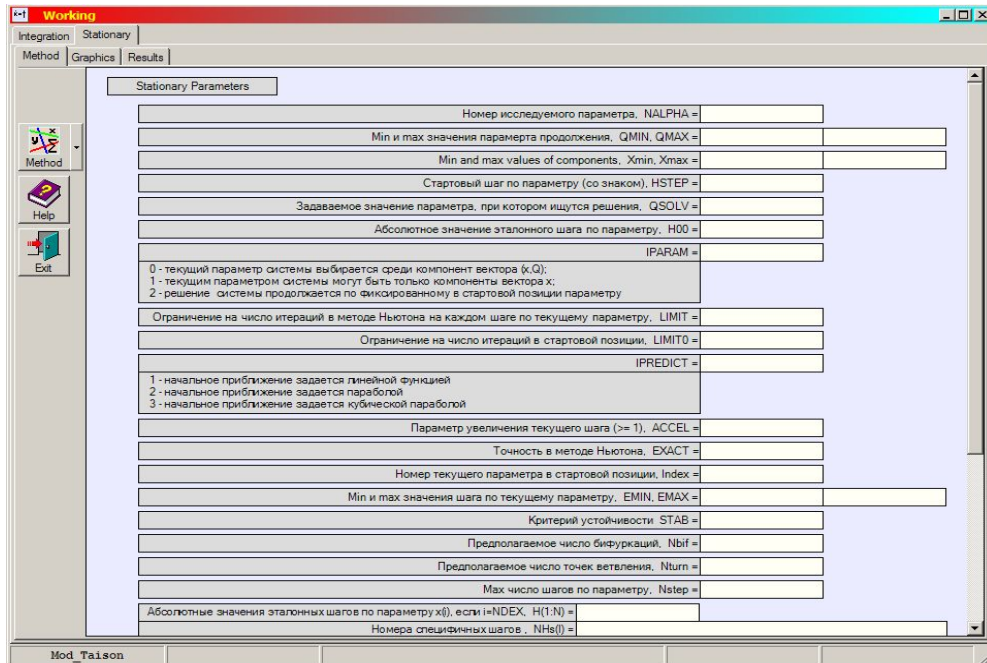



Рисунок 84. Вкладка 'Stationary' формы 'Working'.

Нажмите кнопку 'Method' . Появляется диалоговое окно:

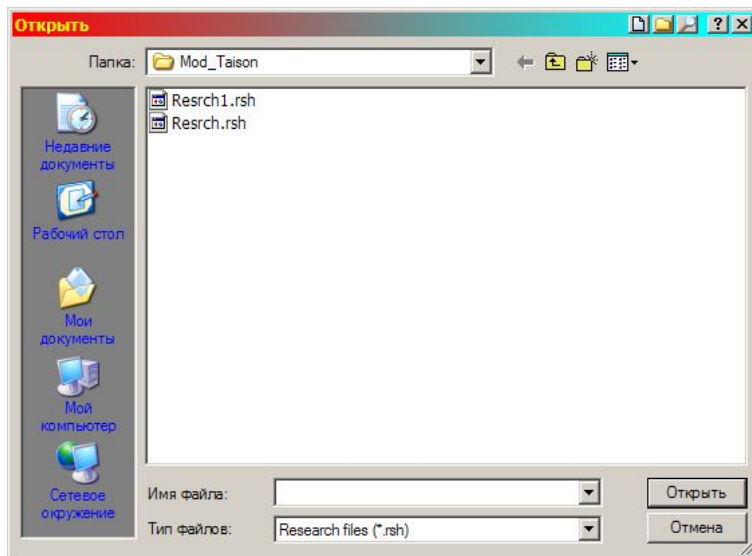


Рисунок 85. Диалоговое окно выбора параметров метода.

Выберите файл, например, 'Resrch.rsh', и нажмите кнопку 'Открыть':

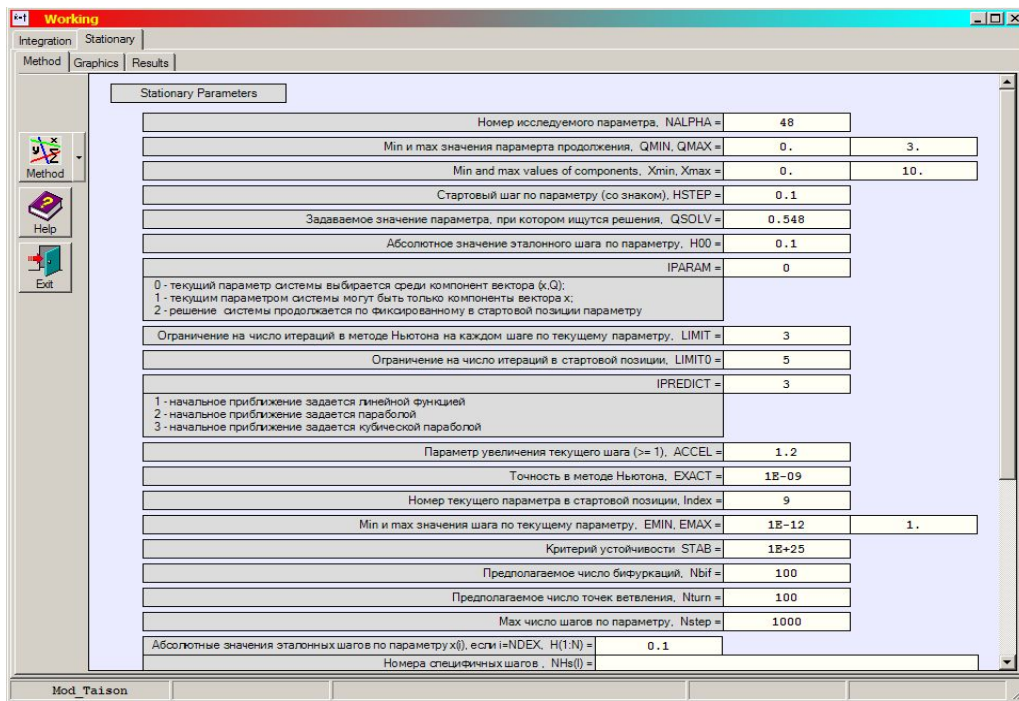


Рисунок 86. Параметры метода из файла 'Resrch.rsh'.

Запуск на расчёт. Визуализация результатов: перейдите на вкладку 'Graphics', нажав на имя вкладки (формы Working):

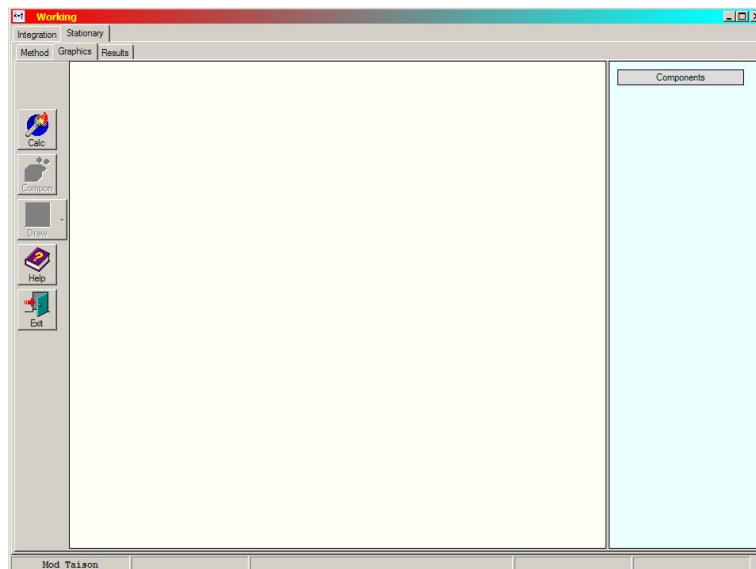




Рисунок 87. Вкладка 'Graphics'.

Нажмите на кнопку 'Calc' . Активируется кнопка 'Compon' . Нажмите на эту кнопку. Активируется кнопка 'Draw' и в панели 'Components' появляется список и цвета выдаваемых компонент:

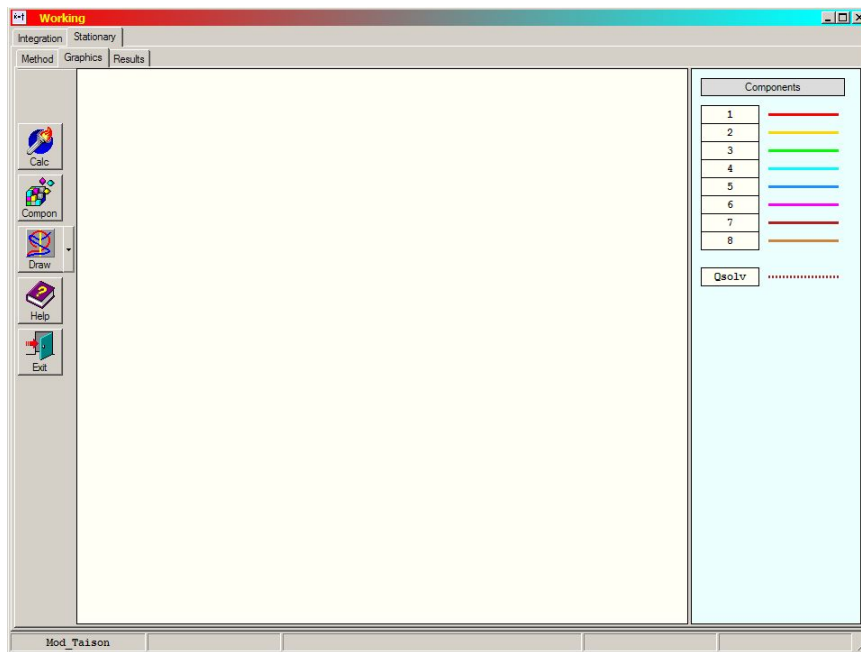



Рисунок 88. Появление в панели 'Components' списка и цветов выдаваемых компонент.

Нажмите на кнопку 'Draw'  . Появляется графики поведения компонент стационарного решения в зависимости от выбранного параметра (в данном случае - $p(48) \sim zMass$):

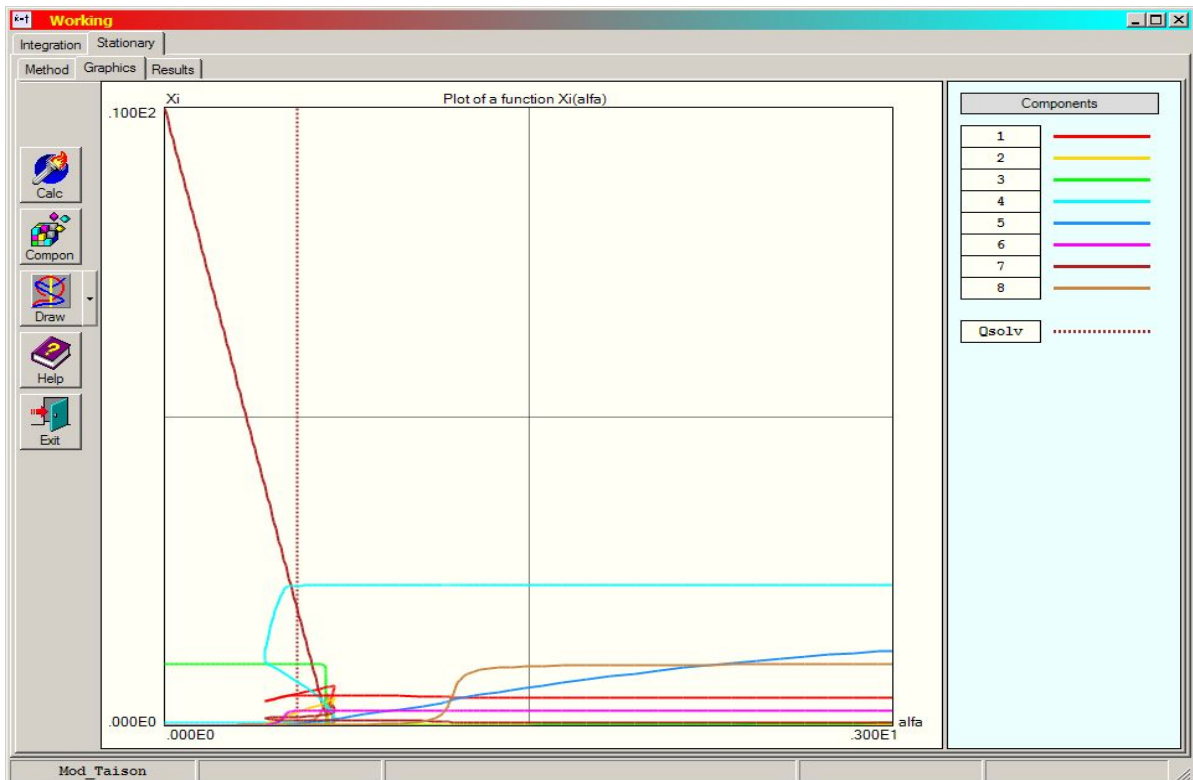



Рисунок 89. Визуализация расчётов модели.

Анализ результатов расчёта: перейдите на вкладку 'Results', нажав на имя вкладки и на кнопку 'Table' . Откроется диалоговое окно со списком файлов результатов:

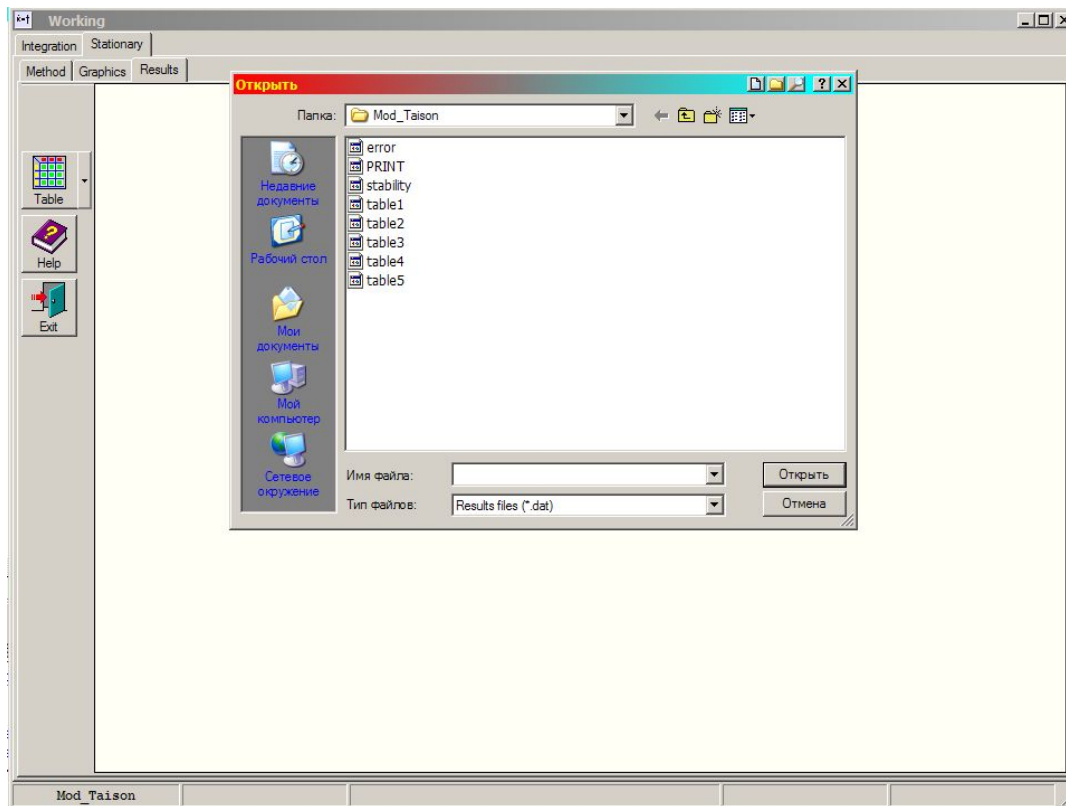


Рисунок 90. Диалоговое окно со списком файлов результатов.

2.1.5. Программная компонента «HGNet»

2.1.5.1. Структура и функциональные возможности программного компонента «HGNet»

Программный компонент «HGNet» предназначен для изучения динамики, предельных состояний (стационарные режимы и предельные циклы), их устойчивости в гипотетических генных сетях (ГГС) различного типа, а также проверки различных гипотез о свойствах ГГС типа (n,k) -критерия.

Понимание текста размещенного в разделе 2.1.5. подразумевает, что пользователь ознакомился со стандартом и форматом файлов, применяемых в программном компоненте «HGNet». Подробное описание стандарта и формата данного файла дано в “ОИОС01”. Подробное описание руководства пользователя дано в “АСНИ-01”.

2.1.5.1.1. Исследование компьютерных моделей в программном компоненте «HGNet»

Запуск HGNet

Запустите исполняемый файл hgnet.exe в директории, в которой размещен HGNet. Появляется диалоговое окно (Рис. 91) следующего вида:

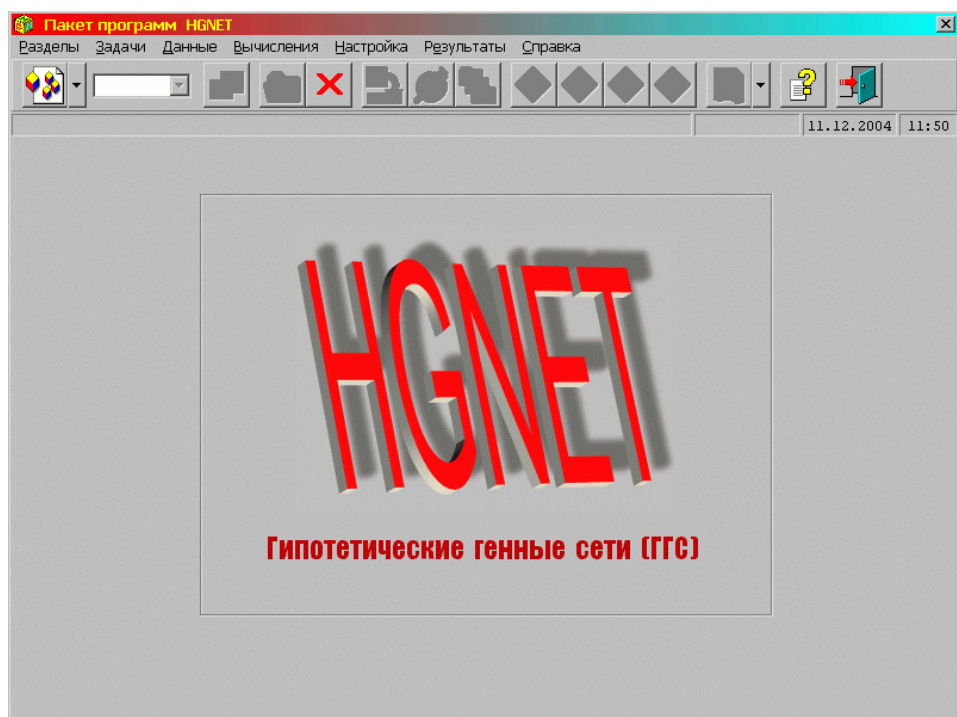


Рисунок 91. Диалоговое окно пакета HGNet.

Выбор раздела

Выберите раздел в соответствующем пункте меню, либо нажмите на кнопку со стрелкой на панели инструментов открывается развёрнутый список всех разделов со всеми подразделами:

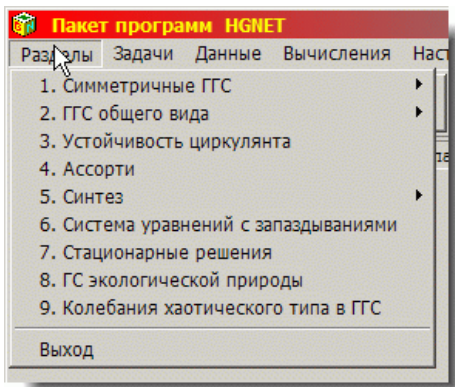


Рисунок 92. Выбор раздела.

Выберите доступный подпункт меню или соответствующую кнопку на панели инструментов: *Выбрать файл и открыть таблицы.*

Выберите директорию задачи для задач с запаздываниями.

Выберите для раздела Ассорти доступный подпункт меню и соответствующую кнопку на панели инструментов: *Ввод исходных данных.*

Появляется название раздела в первой подпанели панели состояния.

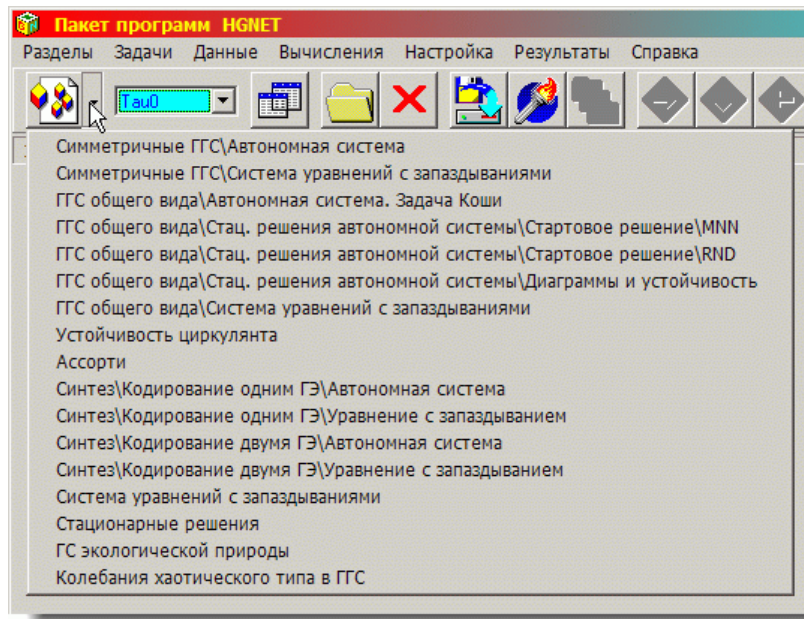


Рисунок 93. Развёрнутый список разделов пакета.

Выбор директории задачи

Выберите директорию задачи из готового списка. После выбора директории становится доступным подпункт: *Выбрать файл и открыть таблицы* - для работы с исходными данными задачи:

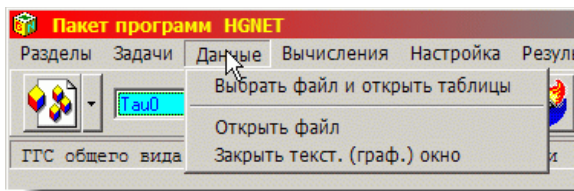


Рисунок 94. Пункт Главного меню Данные.

Задание исходных данных

Выберите файл с именем вида DataX.xxxd, где X - символ (или несколько символов), - для составления банка данных, а xxx - часть расширения, уникальная для каждого раздела. Отредактируете таблицу исходных данных:

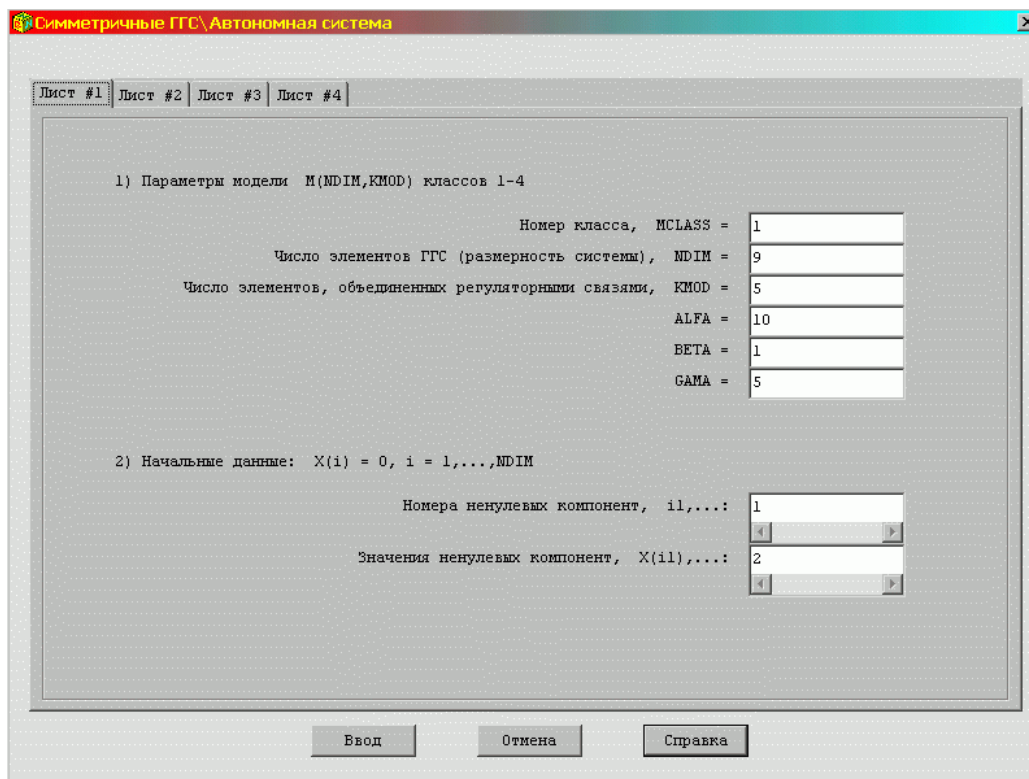


Рисунок 95. Лист исходных данных.

Сохраните отредактированные исходные данные задачи с помощью диалогового окна сохранения файла, открывающегося при нажатии кнопки "Ввод" внизу окна задачи.

Запуск на расчет

Выберите подпункт меню Счёт [и графика] (или нажмите соответствующую кнопку на панели инструментов) начинается собственно счёт.

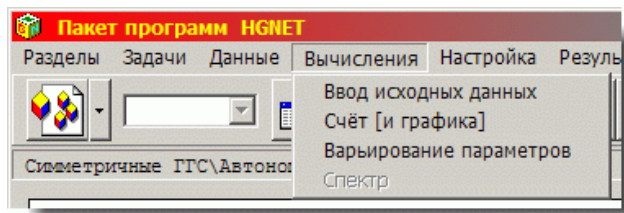


Рисунок 96. Пункт Главного меню Вычисления.

Открывается графическое окно, в котором отображаются графики решений задачи.

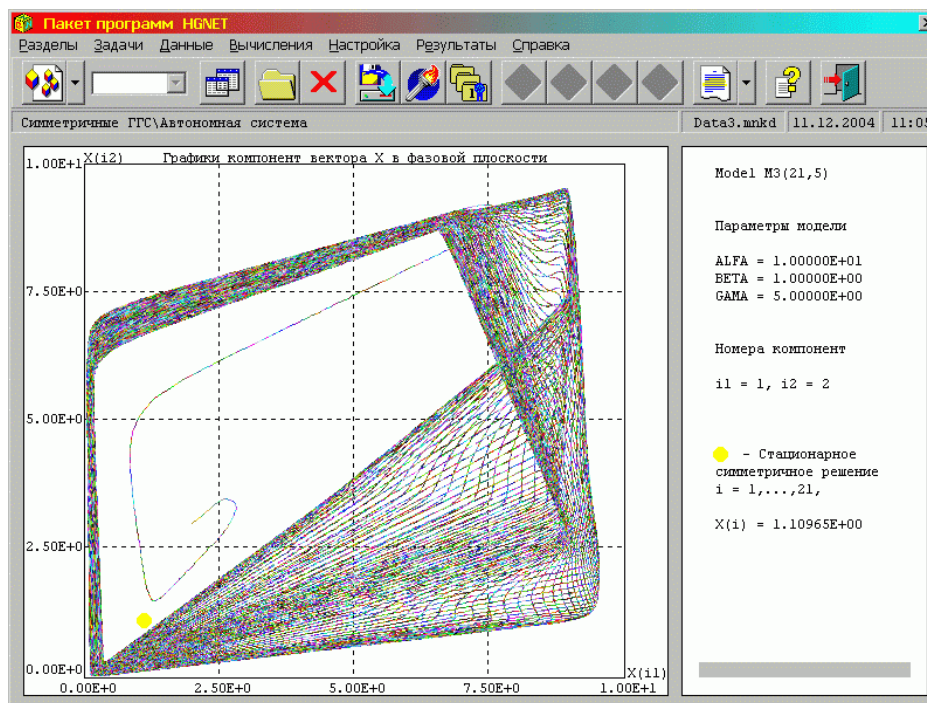


Рис. 2.1.5.- 7. Фазовая плоскость.

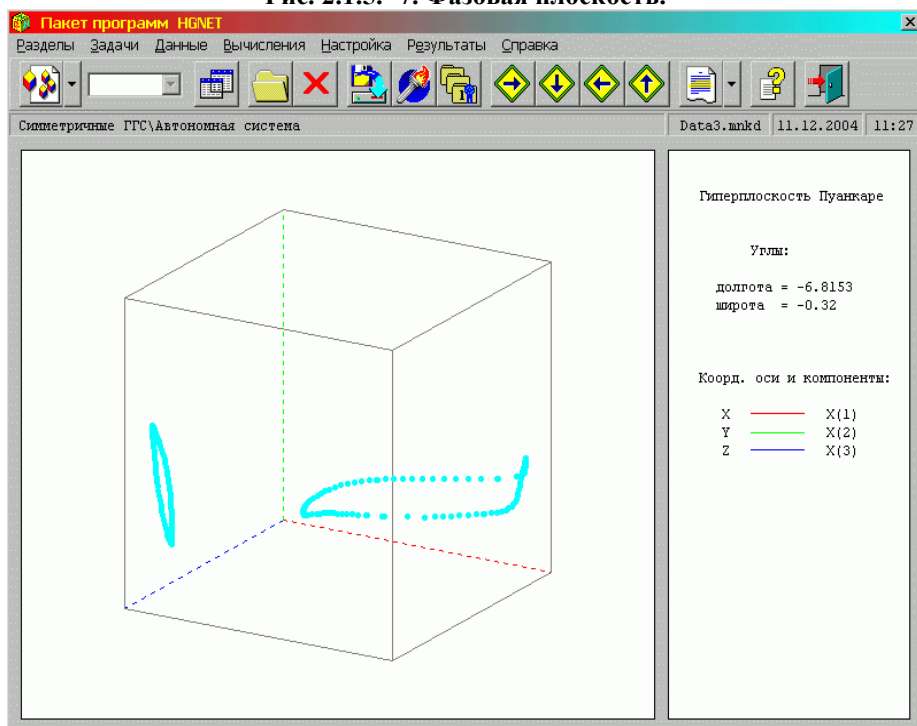


Рисунок 97. Гиперплоскость Пуанкаре.

Опции графики и варьирование параметров

Выберите нужную опцию для рисования графиков.

Отредактируйте значения параметров, приведенных в правом верхнем углу вспомогательного окна.

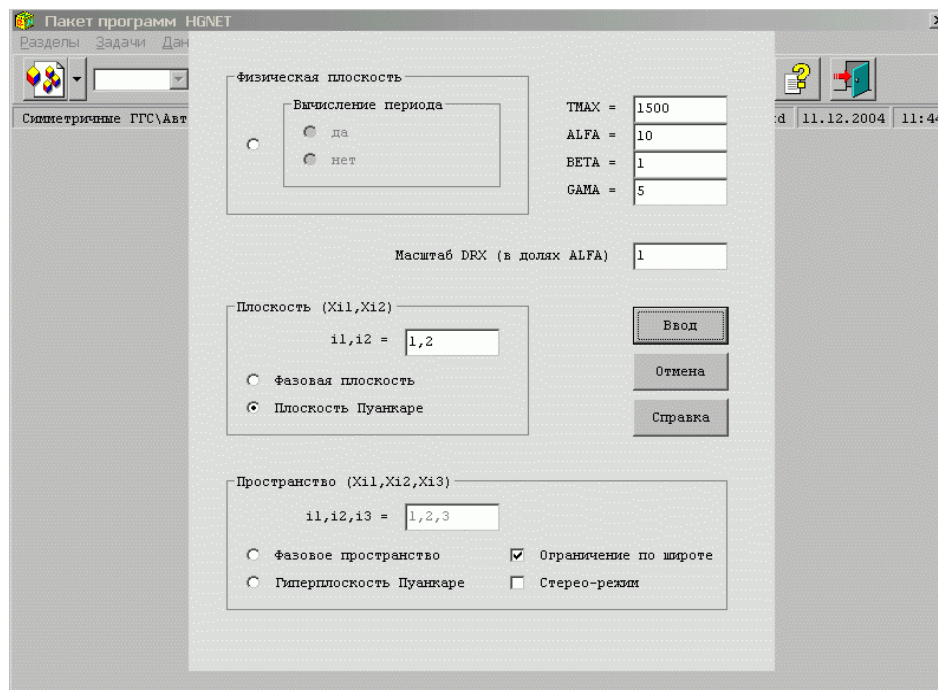


Рисунок 98. Вспомогательное окно варьирования параметров.

Анализ результатов

Выберите подпункт меню (или нажмите на соответствующую кнопку со стрелкой на панели инструментов).

Просмотрите результаты решения, представленные в нескольких файлах:

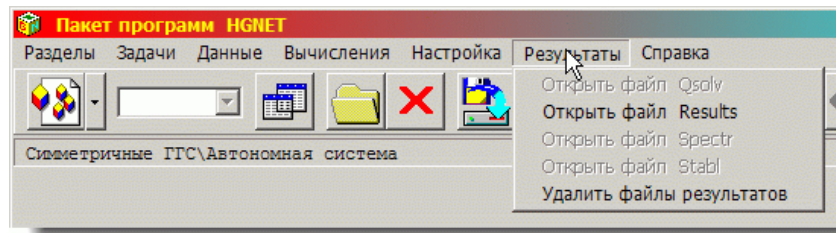


Рисунок 99. Пункт Главного меню Результаты.

```

Модель M1(9,5)
NDIM = 9
KMOD = 5
TMAX = 4.00000E+01
ALFA = 1.00000E+01
BETA = 1.00000E+00
GAMA = 5.00000E+00

Стационарное симметричное решение,
в котором все компоненты равны STAT:
STAT = 1.14169E+00

```

Таблица
зависимостей компонент вектора X от времени t

No	t	X(01)	X(02)	X(03)	X(04)	X(05)
00001	0.00000E+00	2.00000E+00	0.00000E+00	0.00000E+00	0.00000E+00	0.00000E+00
00050	4.90000E-01	1.86035E+00	7.54848E-02	7.61343E-02	7.65833E-02	7.69377E-02
00100	9.90000E-01	1.12965E+00	6.99175E-01	6.97900E-01	6.95027E-01	6.91599E-01
00150	1.49000E+00	6.86579E-01	3.00357E+00	8.76235E-01	8.11483E-01	7.72241E-01
00200	1.99000E+00	4.32689E-01	5.56011E+00	5.35066E-01	4.95789E-01	4.71986E-01
00250	2.49000E+00	4.48115E-01	7.26035E+00	3.24889E-01	3.01067E-01	2.86630E-01
00300	2.99000E+00	1.03403E+00	6.43775E+00	1.97260E-01	1.82811E-01	1.74054E-01
00350	3.49000E+00	6.77267E-01	3.95378E+00	1.21289E-01	1.12526E-01	1.07215E-01
00400	3.99000E+00	4.11777E-01	2.39908E+00	9.33531E-02	8.80413E-02	8.48224E-02
00450	4.49000E+00	2.49965E-01	1.45533E+00	2.79960E-01	2.76764E-01	2.74826E-01

Рисунок 100. Текстовое окно файла результатов решаемой задачи.

2.2. Информационные ресурсы для подсистемы "МОДЕЛИРОВАНИЕ КЛЕТКИ"

Создание математических моделей биологических процессов и систем является одной из важнейших проблем системной биологии. Наличие информационных ресурсов, аккумулирующих экспериментальные данные по различным аспектам функционирования живых систем, представляет собой обязательное условие для компьютерного моделирования. В рамках раздела «МОДЕЛИРОВАНИЕ КЛЕТКИ» для этих целей поддерживаются базы данных GeneNet, KiNET, BiotechPro, GenSensor ConSensor.

2.2.1. База данных GeneNet для прокариот

База данных по метаболическим путям и генным сетям у бактерий является составной частью системы GeneNet (Колчанов и др., 2000; Ananko *et al.*, 2005) и предназначена для компьютерного описания структурно-функциональной организации генных и метаболических сетей прокариот. GeneNet позволяет накапливать разнородные данные по сетям любой сложности, а затем анализировать эти данные компьютерными методами. В базе данных хранится информация о разнообразных молекулярно-биологических объектах, а также о взаимодействиях между этими объектами в рамках сети. Вся информация берется из экспериментальных работ, опубликованных в научных журналах.

Согласно объектно-ориентированному подходу (Schweigert *et al.*, 1995), все компоненты генных сетей разделены на два типа: элементарные структурные компоненты и элементарные взаимодействия (отношения между элементарными структурными компонентами). Каждый из этих типов состоит из нескольких классов объектов, представленных на рис. 101.

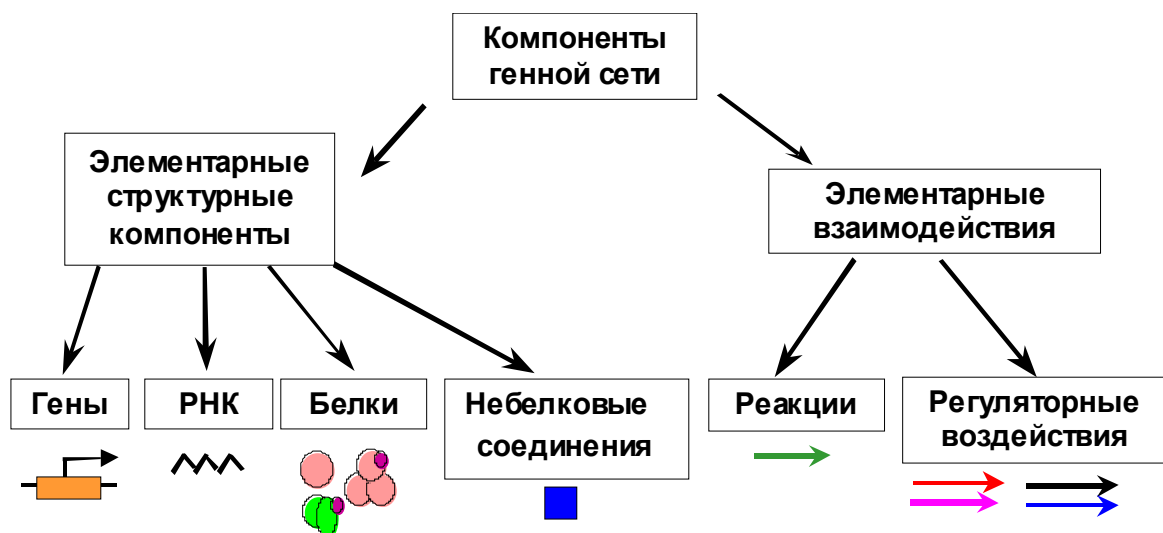


Рисунок. 101. Объекты в генных сетях

База данных содержит описание следующих структурных компонентов генных и метаболических сетей:

- опероны
- гены
- РНК
- белки
- метаболиты

Помимо этого, в базе имеется информация о следующих типах взаимодействия между компонентами сети:

- реакциях
- регуляторных взаимодействиях.

Организация базы данных GeneNet¹

База данных является реляционной и состоит из нескольких таблиц. Количество входов в основных таблицах прокариотической базы данных GeneNet приведено ниже.

Название таблицы GeneNet	Количество входов
GN_OPERON	254
GN_GENE	444
GN_RNA	300
GN_PROTEIN	847
GN_SUBSTANCE	564
GN_RELATION	4329
GN_SCHEME	23
GN_CELL	21

¹ Подробное описание информационных полей и формата таблиц GeneNet приведено в ОИОС_01, раздел 5.3.12.

Все объекты генных сетей делятся на структурные элементы и взаимодействия между ними.

К структурным элементам генной сети относятся: белки, РНК, гены, опероны, метаболиты. Каждый класс структурных элементов описывается в отдельной таблице.

Взаимодействия между компонентами разделяются на реакции и регуляторные воздействия, которые описываются в общей таблице.

При описании белков указывается его имя, синонимы, ссылки на внешние базы данных, код классификатора (только для ферментов), функциональное и мультимерное состояние, субъединичный состав для мультимеров, посттрансляционные модификации и т.д.

Для РНК указывается ген, кодирующий эту РНК, тип РНК (матричная, рибосомальная и т.д.), ссылки на внешние базы данных.

Для гена указывается стандартное название, синонимы, ссылки на внешние базы данных, ссылка на оперон, если он описан в базе данных.

Для оперона указывается стандартное название, синонимы (если есть), ссылки на внешние базы данных, ссылки на гены, входящие в состав данного оперона, если они описаны в базе данных.

Для взаимодействия указывается класс (реакция или регуляторное воздействие), прямое оно или опосредованное, участники взаимодействия (вход и выход), кинетические константы (если известны). Далее для реакций указывается обратимость, а для регуляторных воздействий - тип воздействия (усиливает или подавляет).

Запросы к базе данных

Запросы к базе данных можно осуществлять через редактор генных сетей².

В качестве возможных запросов к базе данных, определяющих требования к системе запросов можно указать следующие:

- Q1: Найти ген, в названии или синонимах которого встречается заданный контекст;
- Q2: Найти белок, в названии или синонимах которого встречается заданный контекст;
- Q3: Найти метаболит, в названии или синонимах которого встречается заданный контекст;
- Q4: Найти аннотированную статью по фамилии одного из авторов;
- Q5: Найти аннотированную статью по фрагменту названия статьи;
- Q6: Найти замкнутые регуляторные циклы на графе сети.

Словари

В рамках прокариотической GeneNet поддерживаются следующие словари:

- названий штаммов бактерий
- названий внешних баз данных, на которые ставятся ссылки в записях GeneNet
- возможных модификаций белка
- названий внешних воздействий, которые могут служить индукторами либо репрессорами для генов, РНК, белков
- названий журналов

² Подробнее см. «Руководство пользователя» АСНИ_01, раздел 4.2.7

Словарь внешних воздействий имеют иерархическую структуру и включают как основные, так и синонимичные названия.

Источники номенклатурных имен объектов базы данных GeneNet

Для прокариотической базы данных GeneNet используются следующие общепринятые зарегистрированные классификаторы:

1. Для ферментов (таблица PROTEINS) в поле 'enzymeClass' проставляется код международного классификатора ферментов.

Пример для входа protein354

enzymeClass 2.7.1.71

2. Для генов *E.coli* (таблица GENES) в поле 'BN' проставляется стандартный номер гена в геноме.

Пример для входа gene204

BN b3281

Источником номенклатурных имен генов и белков *E.coli* являлись базы данных KEGG, EcoCyc

Ссылки на внешние ресурсы

В прокариотической GeneNet, помимо внутренних ссылок между таблицами базы, содержатся также ссылки на внешние базы данных: CAS, ColiBase, EC_TRRD, EcoCyc, EcoGene, EMBL, GenBank, KEGG, PIR, PubMed, RegulonDB, SWISS-PROT, TrEMBL.

Прокариотическая GeneNet связана со следующими информационными ресурсами АСНИ «Системная биология-01»: ProkaTEX, GenBank, EntrezGene, KEGG.

Информационное содержание прокариотической базы данных GeneNet

В ноябре 2006 года база данных содержала описание 20 метаболических сетей *E. coli*. Количество основных компонентов этих сетей приведено ниже.

<i>Название схемы</i>	<i>Число компонентов</i>		
	<i>Оперонов</i>	<i>Белков</i>	<i>Связей</i>
Alanine Biosynthesis	4	10	33
Anaplerotic Reactions	9	23	133
Arginine, Putrescine, and Spermidine Biosynthesis	23	56	205
Aromatic Amino Acids	15	30	130
Aspartate and Asparagine Biosynthesis	7	18	89
Branched Chain Amino Acid Biosynthesis	11	38	124
Cysteine Biosynthesis	6	26	99
Glutamate and Glutamine Biosynthesis	5	34	204
Histidine Biosynthesis	2	21	63
Membrane Transport	3	17	58
Methionine Biosynthesis	9	21	75
One Carbon Metabolism	6	25	68
Pentose Phosphate Pathway	12	35	104
Proline Biosynthesis	5	17	57
Purine Biosynthesis	11	33	126
Pyrimidine Biosynthesis	7	20	61

<i>Название схемы</i>	<i>Число компонентов</i>		
	<i>Оперонов</i>	<i>Белков</i>	<i>Связей</i>
Pyruvate Metabolism	19	52	178
Respiration	22	111	313
Salvage Pathways	30	85	476
Serine and Glycine Biosynthesis	10	32	107

2.2.2. База количественных данных и кинетических характеристик (БД Kinet)

Высокий темп биологических исследований последних десятилетий приводит к появлению большого количества новой информации и вызывает необходимость накопления и обработки этих данных. На сегодняшний день существует большое количество биологических баз данных, содержащих информацию о структуре генов, метаболических путях, ферментативных процессах. В этих базах данных содержится как описательная информация, так и некоторые количественные характеристики биологических объектов и процессов, т.е. константы ферментативных реакций, массы белков, длины нуклеотидных последовательностей. Однако, по крайней мере, в известных базах данных большая группа количественных характеристик еще не затронута – это информация о концентрациях соединений в конкретном типе клеток, органах и организме в целом, а также о динамике этих концентраций. Между тем для математического моделирования необходимо большое количество информации такого типа.

База кинетических данных (КиНЕТ)/Kinetic data database (KiNET) предназначена для аннотации количественных характеристик: информации о константах протекания биохимических реакций, концентрациях соединений и динамике этих концентраций при изменении условий окружающей среды, введении каких-либо веществ в ходе эксперимента, или при изменении, вызванном естественным путем (например, в ходе дифференцировки клетки или при апоптозе). Собираемые данные могут быть получены в различных клеточных типах, тканях, организмах, а также в системах *in vitro*. Информация собирается из опубликованных экспериментальных работ. Базу можно использовать для накопления количественных характеристик, их обработки и анализа. Эта база данных может быть использована при математическом моделировании с помощью автоматического обращения специальных программ к содержащейся информации, что позволит существенно экономить время сортировки данных для конкретной математической модели. Также база будет полезна для более широкого круга исследователей в качестве справочного информационного источника.

Описание работы с базой данных Kinet

Для входа в подсистему «Компьютерных моделей функционирования бактериальных клеток») и с помощью мыши кликнуть надпись «KiNET» в списке операций этой подсистемы. В результате появится HTML-страница (Рис. 102), обеспечивающая доступ к базе данных со списком генных сетей, для которых в базе данных представлена количественная информация.

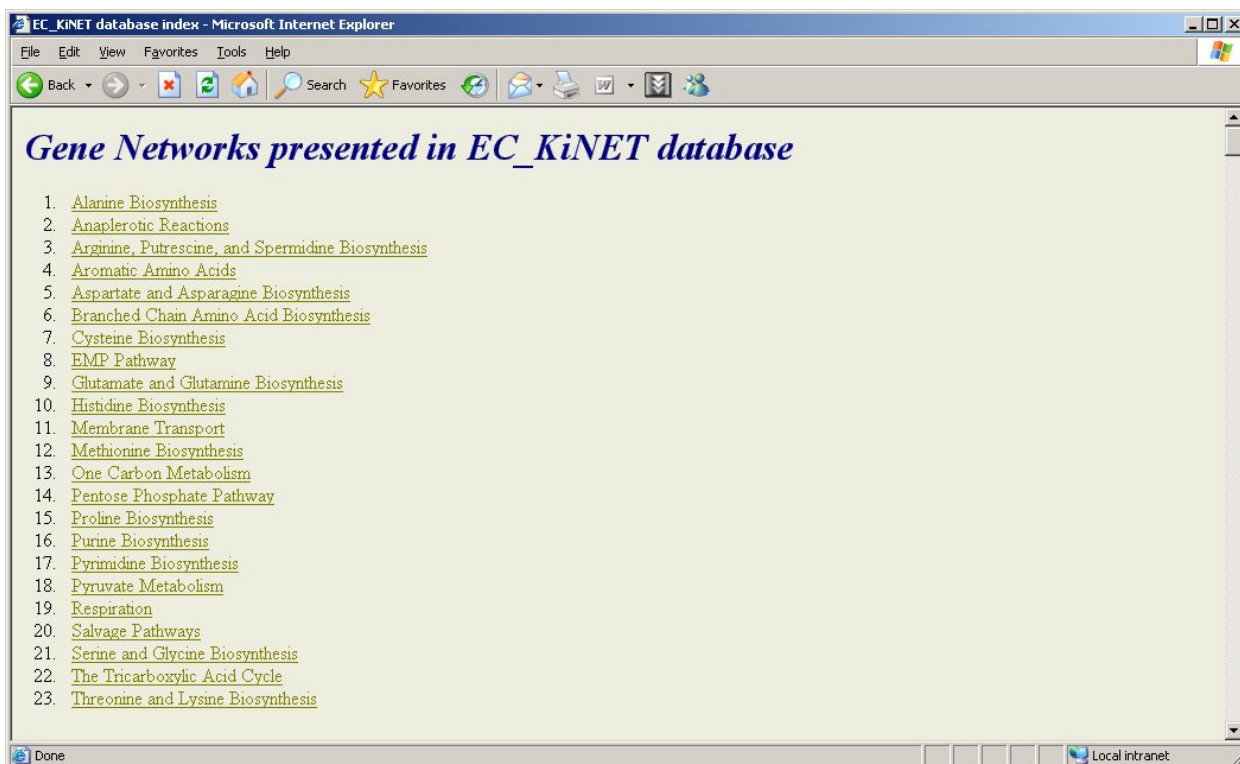


Рисунок 102. HTML-страница базы данных Кинет со списком генных сетей, информация о которых собрана в базе.

Выбрав из представленного списка отдельную генную сеть, можно получить список реакций и молекулярно-генетических процессов, составляющих данную генную сеть (Рис. 103).

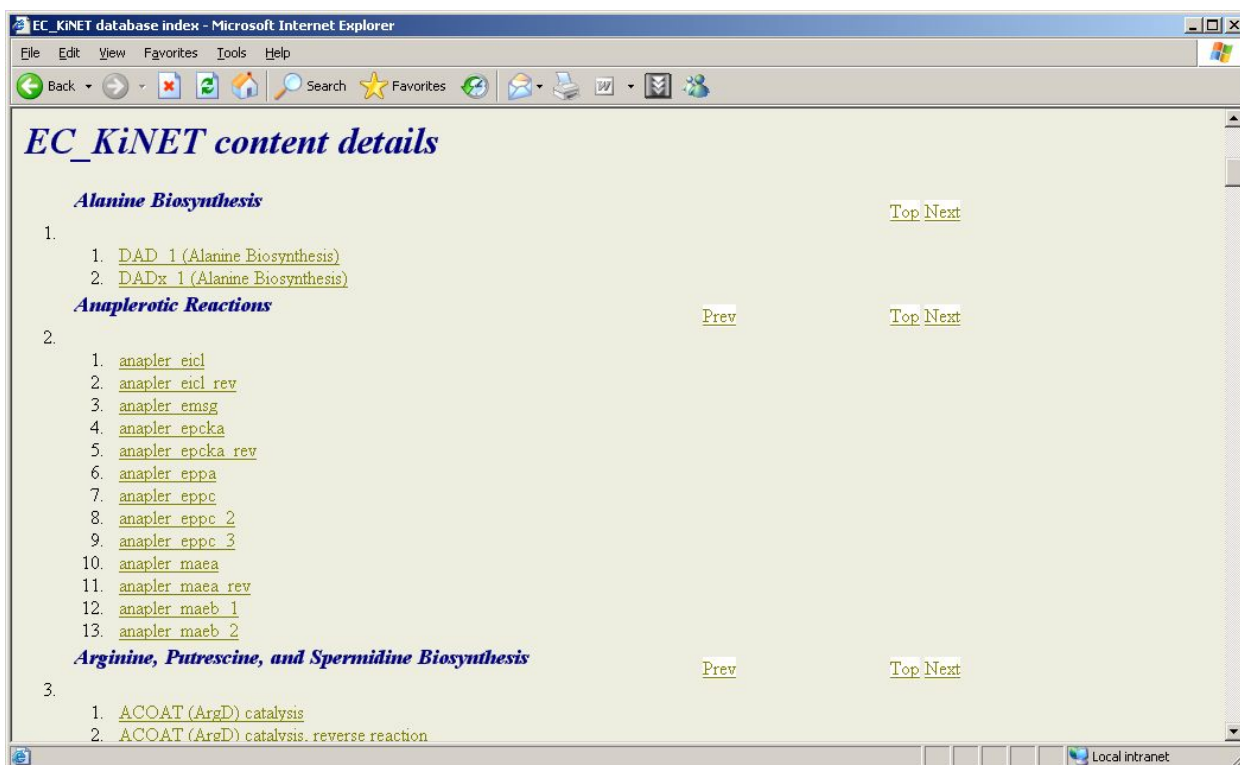


Рисунок 103. Информация о процессах, относящихся к определенной генной сети.

Пользователь может выбрать любую из реакций (или молекулярно-генетический процесс), относящихся к определенной генной сети, и перейдет на страницу с описанием этой реакции (Рис. 104). Описание процесса включает названия веществ-участников (краткое, полное, синонимы), и роли этих веществ в процессе (субстрат, продукт, фермент, индуктор, ингибитор, кофактор).

EC_KINET database entry: Arginine: CPSase catalysis 8 (carAB), GLN-dependent CAP synthesis - Microsoft Internet Explorer

Home Up Prev Next

Reactants and products with their respective roles in the process with symbolic name:

Arginine: CPSase catalysis 8 (carAB), GLN-dependent CAP synthesis

Process specific roles	Name	Synonyms
substrate	Water	
substrate	Carbon dioxide	Carbonic anhydride; Carbonic acid gas;
substrate	Adenosine triphosphate	Adenylypyrophosphate; Adenosine 5'-triphosphate;
substrate	Glutamine	L-2-Aminoglutaramic acid; L-Glutamine; proglumide; Glum; Glumin; Glutamic acid 5-amide;
enzyme	Carbamoyl phosphate synthetase	
product	Carbamoyl phosphate	carbamoyl-P;
product	Glutamate	L-Glutamic acid; 1-amino-propane-1,3-dicarboxylic acid; Glutacid; alpha-aminoglutaric acid; Glt; L-Glutamate; Glut; 2-aminopentanedioic acid; L-glu; L-Glutamic acid;
product	Adenosine diphosphate	Adenosine 5'-diphosphate; Adenosine 5'-pyrophosphate; Adenosine pyrophosphate;
product	Phosphate (inorganic)	Phosphate; Inorganic phosphate; phosphate-inorganic; Orthophosphate;
inducer	Ornithine	(S)-2,5-diaminovaleric acid; (S)-2,5-Diaminopentanoic acid; (S)-2,5-Diaminopentanoate; L-Ornithine;
repressor	Uridine monophosphate	5'Uridylic acid; Uridine 5'-monophosphate; Uridylate; Uridylic acid; Uridine-phosphate;

Рисунок 104. Информация о биохимической реакции или молекулярно-генетическом процессе в базе данных KINET.

Ссылка **Home** направляет на страницу с перечнем генных сетей.

На этой же странице, под описанием участников реакции (Рис. 105) находится описание экспериментов, с помощью которых были получены количественные данные. Описание эксперимента включает три блока:

- описание объекта исследования (**Object**);
 - Species** – краткое название вида.
 - Object status** – статус организма, указывается соответствующее поле, когда исследовался мутантный, трансгенный или нокаутный организм (Wild type/Mutant/Transgen/Knockout).
 - Oxygen status** (unknown, aerobic, anaerobic, microaerobic) – кислородные условия эксперимента, если необходимо.
- описание метода, использованного в эксперименте (**Method**);
 - Name** – краткое название метода, с помощью которого определяли динамику или константы. Например: *colorimetric assay*, *fluorometric assay*, *spectrophotometric assay*. В случае, когда метод в статье не указан, а известно, например, что для определения какой-либо константы использовалось субстратное насыщение, то в это поле вносится *enzyme assay*.

Description – комментарии к полю **Name**. В это поле содержится описание метода измерения констант или динамики концентраций. Также может быть описан метод получения и/или очистки исследуемого вещества.

- условия проведения эксперимента (**Conditions**).

Buffer – информация о растворе, в котором проводилась реакция. Кроме буферной системы раствора в этом поле содержится также информация о других компонентах раствора, не участвующих в биохимической реакции. Например: 50 mM Tris/HCl, 1 mg/ml of bovine serum albumin, 10 mM Mg(Cl)2.

Temperature – значение температуры, при котором проводился эксперимент.

pH – значение pH.

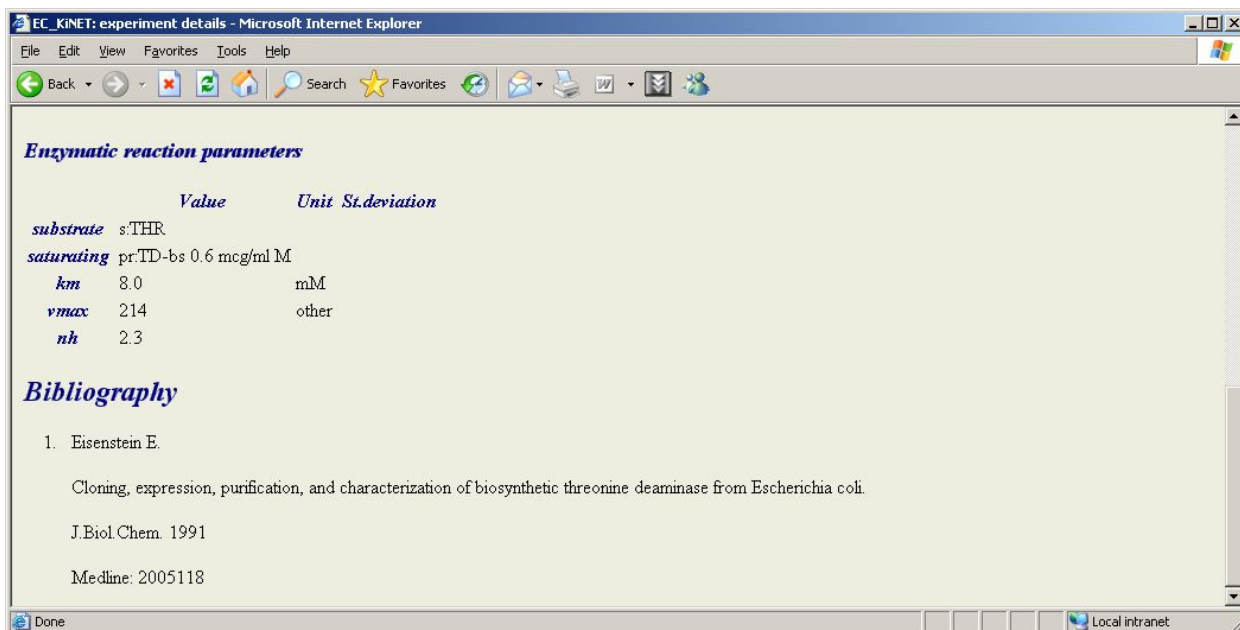
The screenshot shows a web browser window with the title "EC_KINET database entry: Arginine: CPSase catalysis 8 (carAB), GLN-dependent CAP synthesis". The page content is titled "Experiments" and lists five numbered entries. Each entry contains a table with columns for "Object", "Method", and "Conditions".

Entry	Object	Method	Conditions
1. Collected constants ->	wild type		Buffer: 0.1 M Tris-HCl pH 8.0, 0.1 M KCl, 20 mM KHCO ₃ , 10 mM glutamine, 5 mM ATP, 5 mM MgSO ₄
	E. coli	enzyme assay	37Å°C
	Oxygen status: unknown	no data	8pH
2. Collected constants ->	wild type		Buffer: 50 mM HEPES, 100 mM KCl, 40 mM KHCO ₃ , 2 mM ATP, 20 mM MgCl ₂ , 5 mM ORN, 96 units of pyruvate kinase, 20 mM PEP, 200 units of ornithine transcarbamoylase
	E. coli	coupled assay with OTCase	25Å°C
	Oxygen status: unknown	Measuring the amount of citrulline formed in a coupled assay with ornithine transcarbamoylase and ORN	7.6pH
3. Collected constants ->	wild type		Buffer: 0.1 M Tris-HCl, pH 8.0, 0.1 M KCl, 20 mM NaHCO ₃ , 10 mM glutamine, 7 mM MgSO ₄ , and 5 mM ATP
	E. coli	enzyme assay	37Å°C
	Oxygen status: unknown	no data	8pH
4. Collected constants ->	wild type		Buffer: 50 mM HEPES, 100 mM KCl, 10 mM ATP, 20 mM MgCl ₂ , 20 mM NaHCO ₃ , 10 mM ORN, 10 mM GLN, 0.8 unit of ornithine transcarbamoylase
	E. coli	coupled assay with OTCase	25Å°C
	Oxygen status: unknown	Measuring the amount of citrulline formed in a coupled assay with ornithine transcarbamoylase and ORN	7.6pH
5. Collected constants ->	wild type		Buffer: 0.1 M Tris-HCl, pH 8, 0.1 M KCl, 10 mM glutamine, equimolar ATP and MgCl ₂ (10 mM), 20 mM NaHCO ₃
	E. coli	coupled assay with OTCase	37Å°C

Рисунок 105. Описание условий проведения эксперимента для получения количественных данных.

Каждому описанию эксперимента соответствует гиперссылка **Collected constants** для перехода на HTML-страницу с описанием констант, полученных в ходе эксперимента. На этой странице указываются значения и единицы измерения, полученных в ходе эксперимента

констант (Рис. 106). Информация о библиографическом источнике **Bibliography** находится под списком констант (Рис. 106).



The screenshot shows a Microsoft Internet Explorer window titled "EC_KINET: experiment details". The main content is divided into two sections:

Enzymatic reaction parameters

	Value	Unit	St.deviation
substrate	s:THR		
saturating	pr:TD-bs 0.6 mcg/ml	M	
km	8.0	mM	
vmax	214	other	
nh	2.3		

Bibliography

1. Eisenstein E.
Cloning, expression, purification, and characterization of biosynthetic threonine deaminase from Escherichia coli.
J.Biol.Chem. 1991
Medline: 2005118

Рисунок 106. Информация о константах протекания реакций, описанных в базе данных KiNET.

2.2.3. База данных BiotechPro. Решение задач бактериальной метаболической инженерии.

Значительная часть продуктов, образующихся в ходе метаболизма бактериальной клетки, обладает физиологической активностью и представляет собой практическую ценность для различных отраслей промышленности. Современный уровень развития молекулярной биологии и технологии рекомбинантных ДНК позволяют увеличить пул генов и уровень экспрессии бактериального продукта в штамме в желаемом направлении (Stephanopoulos, 2002). Для реализации таких задач необходима информационная поддержка средствами компьютерного анализа экспериментальных данных и компьютерного моделирования процессов. Информация о разного рода продуктах, синтезируемых различными видами бактерий, штаммах-продуцентах, а также молекулярно-генетических системах, контролирующих синтез, разбросана во множестве статей, каждая из которых содержит результаты исследований лишь отдельных аспектов синтеза целевого продукта.

Первым этапом такой работы является создание специализированной базы данных, накапливающей информацию о важнейших биотехнологически важных продуктах, синтезируемых микроорганизмами, путях его синтеза и штаммах продуцентах, используемых в различных областях биотехнологии. Существующие базы данных не способны в полном объеме отвечать этим требованиям. Первичным навигатором в этой области является база данных "The Prokaryotes" - электронный ресурс, который может быть использован в качестве внешнего источника данных о важнейших биотехнологически значимых микроорганизмах-продуцентах. Данная база содержит информацию в основном, о промышленных штаммах продуцентах, уже нашедших широкое применение в биотехнологических процессах. Базы данных BSD: the Biodegradative Strain Database и Biocatalysis/Biodegradation Database – это

узко специализированные базы данных, содержащие информацию о штаммах- продуцентах ферментов, утилизирующих опасные химические вещества (Urbance *et al.*, 2003, Ellis *et al.*, 2006)

В базе данных **BiotechPro** интегрирована информация о наиболее известных к настоящему времени микроорганизмах, метаболиты которых используются или могут использоваться в биотехнологических процессах. Данные в базу заносятся на основе аннотации научных публикаций

Базовым программным средством, обеспечивающим поиск и навигацию по базе данных **BiotechPro**, а также ее интеграцию с внешними информационными и программными ресурсами является Sequence Retrieval System (SRS). Информация, содержащаяся в базе **BiotechPro** в SRS представлении распределена по 2 таблицам: BiotechProduct и BiotechStrain.

BiotechProduct содержит описание биотехнологически значимого продуктов, продуценты которых описаны в базе BiotechStrain.

BiotechStrain содержит описание штаммов продуцентов биотехнологически значимого продуктов. Таблицы связаны между собой гиперссылками.

В базе данных **BiotechPro** имеются ссылки на внешние базы данных PubMed, ProkaTEX, KiNET и GeneNet.

Описание работы с BiotechPro

База данных **BiotechPro** предназначена для накопления информации о биотехнологически значимых продуктах и их продуцентах.

Поиск информации в базе данных **BiotechPro** возможен как по продуктам, так и по продуцентам. Для этого пользователь должен выбрать одну из 2 таблиц базы данных: BiotechProduct и BiotechStrain (Рис. 107).

SELEX DE	116 21-Apr-2006 Selex Databases
SELEX BIB	70 21-Apr-2006 Selex Databases
SELEX TOOLS	43 21-Apr-2006 Selex Databases
ASPD ALIGN	196 21-Apr-2006 Artificial selected proteins/peptides databas
ASPD REF	112 21-Apr-2006 Artificial selected proteins/peptides databas
SYSTEM	148 21-Apr-2006 Experimental Systems
CROSS TEST	51 21-Apr-2006 Experimental Systems
rSNP DB	46 21-Apr-2006 rSNP_Guide
rSNP BIB	27 21-Apr-2006 rSNP_Guide
TRSIG EXP	106 21-Apr-2006 Translation Signals
TRSIG LONG	129 21-Apr-2006 Translation Signals
TRSIG OBJ	38 21-Apr-2006 Translation Signals
TRSIG SEQ	225 21-Apr-2006 Translation Signals
TRSIG PRODUCT	32 21-Apr-2006 Translation Signals
PROF PAT ALN	138787 10-Jul-2006 Some Proteins databases
PROF PAT BNK	138785 10-Jul-2006 Some Proteins databases
NUCLEOSOME	112 21-Apr-2006 Nucleosome databases
TGP GENE	34 26-Jul-2006 TransGene Promoter Database
TGP PROMOTER	77 26-Jul-2006 TransGene Promoter Database
TGP SEQUENCE	75 26-Jul-2006 TransGene Promoter Database
BIOTECH PRODUCT	66 15-Jul-2006 BiotechPro Database
BIOTECH STRAIN	138 15-Jul-2006 BiotechPro Database

SRS 6.1.3.11 | [feedback](#)

Рисунок 107. Таблицы базы данных: BiotechProduct и BiotechStrain.

Поиск по таблице BiotechProduct

Например, при поиске продукта, необходимо выбрать таблицу BiotechProduct, после чего, загружается интерфейс полей карточки BiotechProduct и выбирается кнопка «Search» (Рис. 108).

Name	Short Name	Type	No of Keys	No of Entry References	Indexing Date	Status
ProductID	pid	id	66	66	15-Jul-2006	ok
Product	prd	index	208	269	15-Jul-2006	ok
Date	dat	num	45	66	15-Jul-2006	ok
Author	ath	show	0	0		not indexed
Species	spc	index	95	197	15-Jul-2006	ok
Pathway	pwy	index	216	427	15-Jul-2006	ok
Comment	cmm	index	846	1881	15-Jul-2006	ok
Application	apl	index	176	386	15-Jul-2006	ok
Keyword	kwd	index	207	387	15-Jul-2006	ok
StrainID	str	index	132	132	15-Jul-2006	ok
LinkGeneNet	lgn	show	0	0		not indexed
LinkProkaTEX	lpc	show	0	0		not indexed
LinkRaNet	lra	show	0	0		not indexed

Рисунок 108. Интерфейс полей карточки BiotechProduct.

После описанной процедуры появляется страница с описанием индексированных полей карточки (Рис. 109).

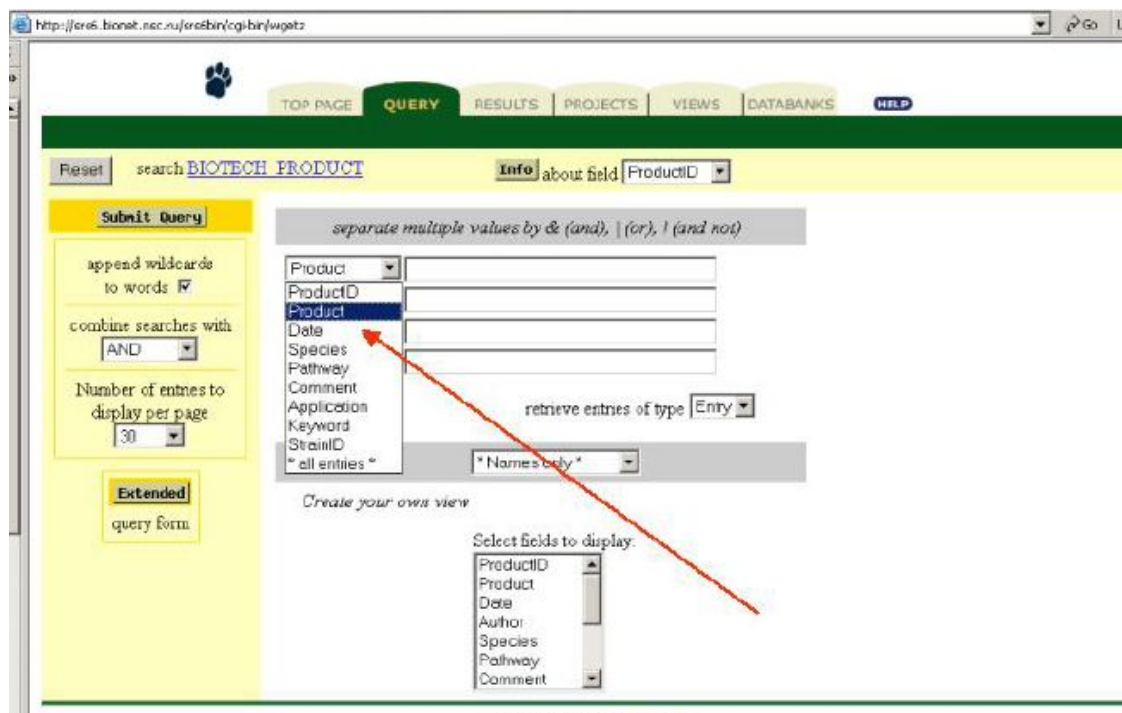


Рисунок 109. Страница с описанием индексированных полей карточки.

Для осуществления поиска, пользователю необходимо ввести название продукта и/или название вида-производителя и нажать кнопку «SUBMIT QUERY». В приведенном на рис. 110 на примере предлагается найти биотехнологически значимый продукт тагатозу, синтезируемую E.coli.

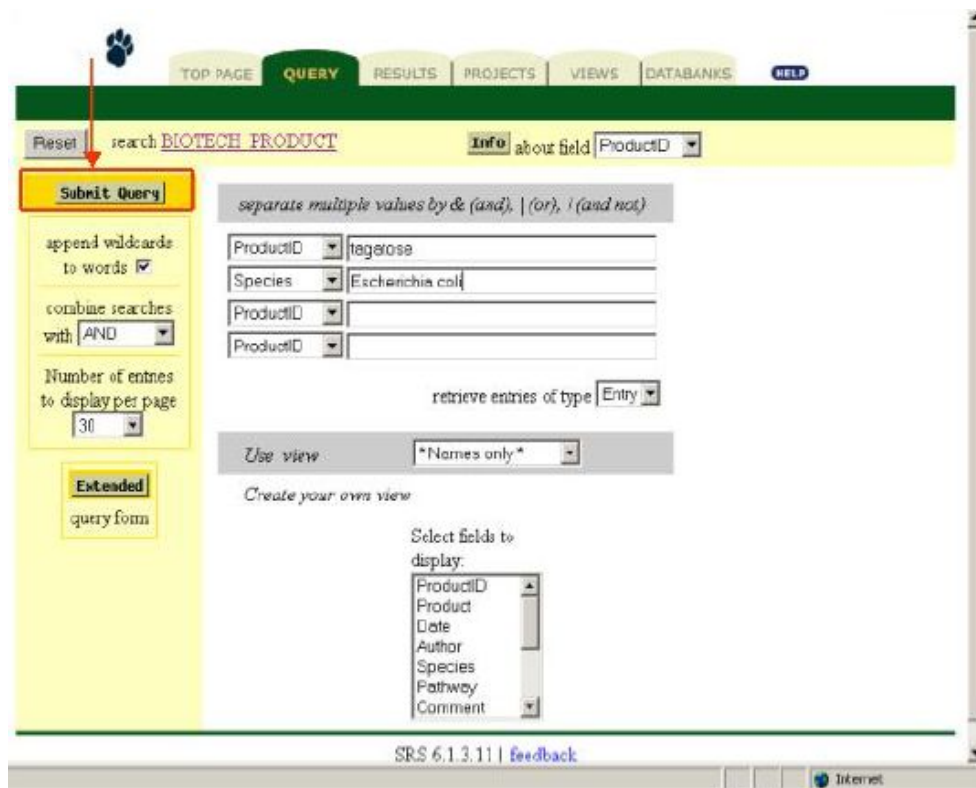


Рисунок 110. Поля поиска.

Результат поиска представлечн на рисунке 111.



Рисунок 111. Результаты поиска.

Нажав на ссылку, пользователь получит информацию о тагатозе, включая информацию и о ее продуцентах (Рис. 112).

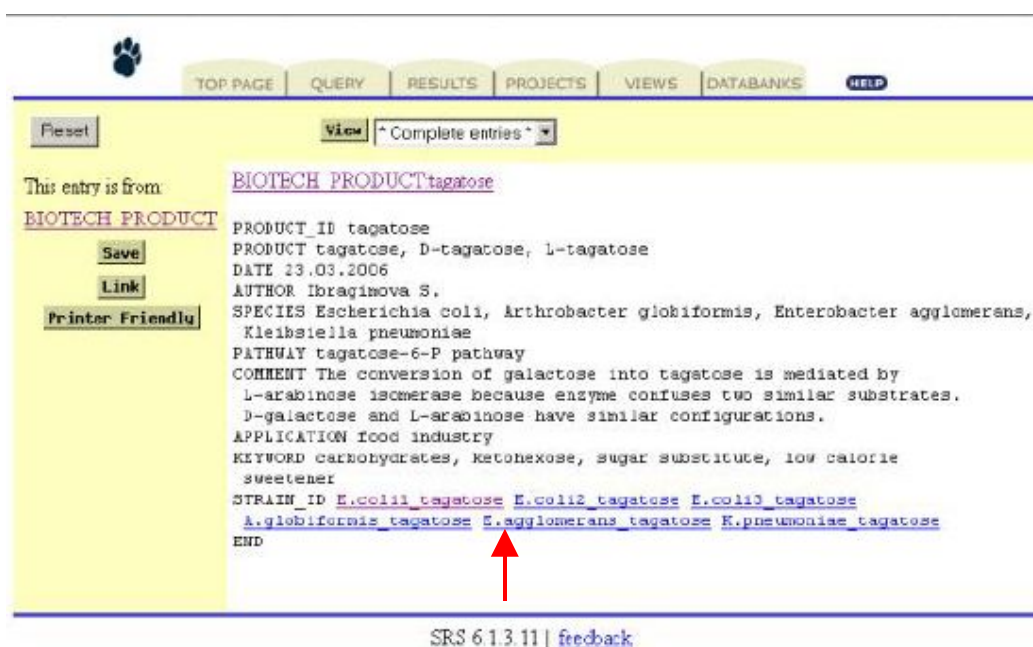


Рисунок 112. Результаты запроса.

Описание отдельного продуцента можно получить, нажав на интересующий Вас вид микроорганизма в поле STRAIN_ID.

В результате пользователь получит карточку интересующего штамма-продуцента, содержащую информацию о количестве синтезируемого продукта и используемом субстрате (Рис. 113).

TOP PAGE QUERY RESULTS PROJECTS VIEWS DATABANKS HELP

Reset View Complete entries

This entry is from: [BIOTECH STRAIN E.coli2 tagatose](#)

BIOTECH STRAIN

STRAIN_ID E.coli2_tagatose

STRAIN E.coli MC1061 (p6xHThAI119)

E.coli strain transformed with plasmid (p6xHThAI119) expressing thaI gene (corresponding to L-arabinose isomerase) of Thermus sp. IM6501 under the control of Bacillus licheniformis maltogenic amylase promoter and tagged by six histidines at its amino terminus for easy purification.

PRODUCTIVITY D-tagatose production by the purified ThaI was carried out with 0.1% galactose solution as a substrate. A 54% conversion yield of D-galactose to D-tagatose was obtained at 60 °C over 3 days.

PRODUCT ID tagatose

REFERENCE Kim J.W., Kim Y.W., Roh H.J., Kim H.Y., Cha J.H., Park E.E., Park C.S. Production of tagatose by a recombinant thermostable L-arabinose isomerase from Thermus sp. IM6501. Biotechnol Lett., 2005, 25, 963-967.

PUBMED [12609032](#)

END

SRS 6.1.3.11 | [feedback](#)

Рисунок 113. Карточка интересующего штамма-продуцента.

Получить информацию о штаммах продуцентах данного продукта у других микроорганизмов пользователь может, нажав на название продукта в поле PRODUCT_ID, вернуться в карточку продукта, и получить информацию о каждом штамме продуцента, последовательно нажимая на название микроорганизма левой клавишей мыши в поле STRAIN_ID.

Примеры типовых запросов для таблицы BiotechProduct:

Найти микроорганизмы, синтезирующие определенный продукт (поиск по полю PRODUCT);

Найти организмы, продукты, которых можно использовать в определенных биотехнологических процессах (поиск по полю APPLICATION);

Выяснить, способен ли интересующий Вас микроорганизм синтезировать данный продукт

(поиск по полям SPECIES и PRODUCT);

Выяснить есть ли информация в базе о конкретном химическом классе продуктов (поиск по полю KEYWORD).

Ссылки на внешние базы данных PubMed, ProkaTEX, KiNET и GeneNet обеспечивают возможность получения дополнительной информации об интересующем продукте, если такая информация в соответствующих базах данных имеется:

Из БД GeneNet о путях синтеза данного продукта;


Из БД ProkaTEX о регуляции транскрипции генов участвующих в синтезе продукта;

Из БД KiNet о константах и динамических переменных рассматриваемого процесса.

Поиск по таблице BiotechStrain

При поиске штамма-продуцента необходимо выбрать таблицу **BiotechStrain** и загрузить интерфейс полей карточки BiotechStrain. Для осуществления поиска в этой таблице нажать на кнопку «Search»(Рис. 114).

[TOP PAGE](#) | [QUERY](#) | [RESULTS](#) | [PROJECTS](#) | [VIEWS](#) | [DATABANKS](#) | [HELP](#)



Name **BIOTECH_STRAIN**

Status The current release has 135 entries and was indexed 15-Jul-2006.

Description This database accumulates information about strains that produce industrially important products.

Data fields in SRS

Name	Short Name	Type	No of Keys	No of Entry References	Indexing Date	Status
StrainID	sid	id	135	138	15-Jul-2006	ok
Strain	str	index	714	1890	15-Jul-2006	ok
Productivity	prd	index	999	3768	15-Jul-2006	ok
ProductID	pid	index	68	138	15-Jul-2006	ok
Reference	ref	index	880	2689	15-Jul-2006	ok
Pubmed	pbm	show	0	0		not indexed

Links

From Databank	Entries Linked	To Databank	Entries Linked	Links Total	Indexing Date
BIOTECH PRODUCT	63	BIOTECH STRAIN	132	132	15-Jul-2006
BIOTECH STRAIN	132	BIOTECH PRODUCT	63	132	15-Jul-2006

SRS Description [Structure](#) ([srsgen.i](#), [srsdb.i](#), [href.i](#))
[Syntax](#)
[Information](#)

SRS 6.1.3.11 | [feedback](#)

Рисунок 114. Таблица BiotechStrain.

Для осуществления поиска, пользователю необходимо выбрать одну или две строки из 4 поисковых строк системы SRS, ввести поисковый термин(ы) и нажать кнопку «SUBMIT QUERY». На рис. 115 приведен пример поиска штаммов-продуцентов тагатозы.

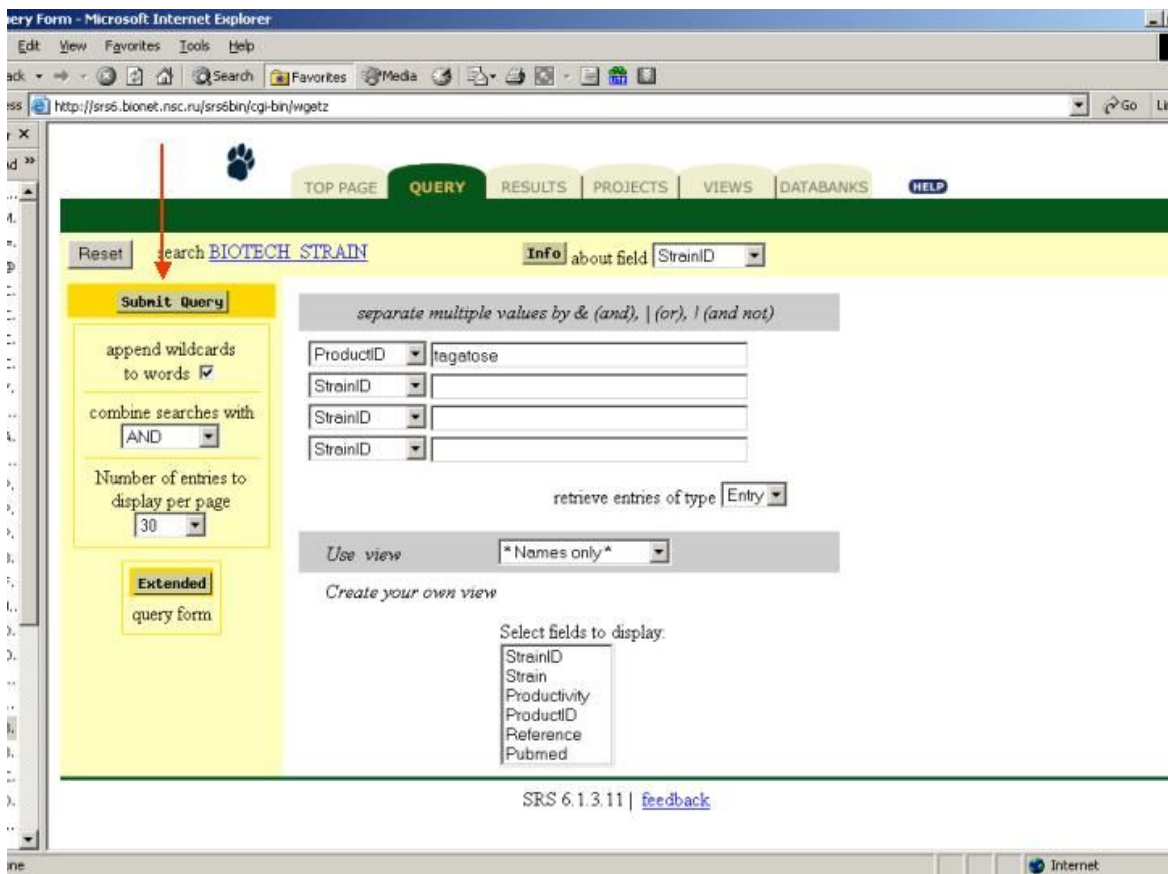


Рисунок 114. Пример поиска штаммов-продуцентов тагатозы.

В результате поиска будет получен список штаммов-продуцентов тагатозы (Рис. 115).

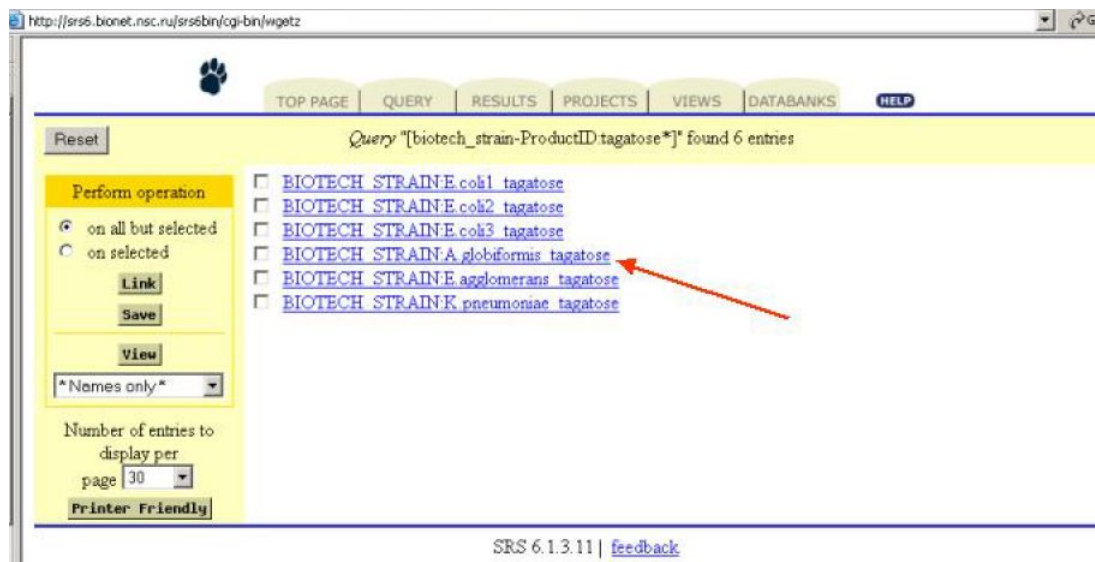


Рисунок 115. Результат поиска.

Нажав левой клавишей мыши на каждый штамм, пользователь получит доступ к карточке штамма.



Рисунок 116. Карточка штамма.

Данные о продукте (пути его биосинтеза, области применения) можно получить, нажав левой клавишей мыши на название продукта в поле PRODUCT_ID.

Примеры типовых запросов для таблицы BiotechStrain:

Найти штаммы-продуценты интересующего Вас продукта (поиск по полю PRODUCT_ID);

Найти штаммы – продуценты, утилизирующие определенный субстрат (поиск по полю PRODUCTIVITY, KEYWORD);

Выяснить, способен ли интересующий Вас микроорганизм синтезировать данный продукт (поиск по полям SPECIES, PRODUCT).

Полученная пользователем информация может быть сохранена на компьютере пользователя с помощью стандартных функций интернет браузера в виде текстового или html-файла и использована непосредственно для работы или же для планирования генно-инженерного эксперимента.

2.2.4. Базы данных GenSensor и ConSensor для дизайна геносенсорных конструкций

Геносенсоры – это искусственные генетические конструкции, которые при введении в клетку способны активировать свою экспрессию в ответ на токсическое воздействие внешних факторов. Чувствительным звеном таких конструкций являются промоторы генов, участвующих в защите клетки от повреждающих воздействий (Zheng M. et al., 2001; Khil P.P. and Camerini-Otero R.D. 2002; Au N. et al., 2005 и др.). Они связаны с репортерным геном, появление продукта которого свидетельствует об активации системы. В качестве генов-эффекторов используются гены, кодирующие структуру люминесцентных и флюоресцентных белков бактерий (*Vibrio fischeri* или *Photobacterium luminescens*), светлячка (*Photinus pyralis*) медузы (*Aequorea victoria*) и кораллов (*Dictyosoma coral*), которые легко детектируются (Nakkila K. et al., 2002).

Разработка и создание геносенсорных конструкций важно с разных точек зрения. Во-первых, они имеют большое прикладное значение как индикаторы самых разных аспектов все более ухудшающихся условий внешней среды, а во-вторых, позволяют провести

фундаментальные, экспериментальные исследования устойчивости и адаптации биологической системы к конкретным воздействиям на уровне отдельной клетки. Это, с одной стороны, экспериментальная проверка результатов моделирования, а с другой – стимуляция развития аппарата моделирования без которого невозможен прогресс в понимании поведения сложных биологических систем.

Успех работы по разработке и созданию высокочувствительных геносенсорных систем во многом зависит от потенциальных возможностей гена-сенсора, т.е. от чувствительности и уровня индукции гена-кандидата к различным внешним воздействиям. Поиск такого гена-кандидата значительно облегчается, если существуют базы данных, в которых подобная информация накапливается. К сожалению, в существующих к настоящему времени базах данных по регуляции транскрипции генов прокариот (EcoCyc, RegulonDB, DPInteract и др.), данных о чувствительности генов к внешним воздействиям нет (Keseler I.M., et al., 2005; Salgado H. et al., 2006; Robison K. et al., 1998). Поэтому первым и необходимым этапом в развитии работ по созданию геносенсорных систем является создание специализированной базы данных, где эта информация будет накапливаться. Это значительный резерв при развитии работ в любом направлении, который не потеряет своего значения и в дальнейшем. Учитывая все вышеперечисленные обстоятельства, мы создали базу GenSensor, которая является для нас источником информации, позволяющей вести поиск генов-кандидатов, удовлетворяющих конкретным требованиям.

Не менее важным является наличие информации об уже существующих конструкторских разработках. Корректная оценка новизны, чувствительности и эффективности созданной сенсорной конструкции возможна только в свете уже существующих данных. Тем более что число описанных в литературе различных геносенсоров с каждым годом становится все больше и для комплексной оценки их эффективности и сравнения между собой требуется систематизация этих данных. Учитывая все вышеперечисленное, мы разработали формат и структуру базы данных ConSensor, в которой идет накопление сведений об уже существующих, описанных в научной литературе геносенсорных конструкциях, включая описание всех компонентов биосенсорной системы и ее чувствительность к внешним воздействиям.

База данных GenSensor.

База данных GenSensor предназначена для накопления информации, необходимой для создания геносенсорных конструкций на основе бактериальных генов и содержит данные о структуре бактериальных промоторов, экспрессия которых активируется в ответ на определенное внешнее воздействие, механизмах индукции, а также условий, при которых наблюдается максимальный ответ на данный тип воздействия.

Отличительной особенностью базы данных GenSensor является то, что она содержит только экспериментально подтвержденную информацию. Источником информации для базы является аннотация научных публикаций. Единицей входа в базу является ген и его описание включает: (1) информацию, идентифицирующую ген (название гена, синонимы, название оперона, в структуру которого он входит, положение гена на хромосоме, уникальный номер гена в геноме бактерии, ссылки на базу GenBank/EMBL); (2) описание структуры промоторов гена (границы промотора, старты транскрипции; последовательности промоторов; названия транскрипционных факторов (ТФ), взаимодействующих с промотором, и их влияние на экспрессию гена); (3) индукторы, активирующие промотор и условия, в которых промотор активируется; (4) ТФ, промоторы и сайты, ответственные за активацию или дерепрессию гена тем или иным индуктором; (5) описание ТФ, связывающихся с промотором (названия, синонимы, видовое происхождение, активная форма).

Формат описания различных компонентов базы GenSensor приведен на рис. 117.

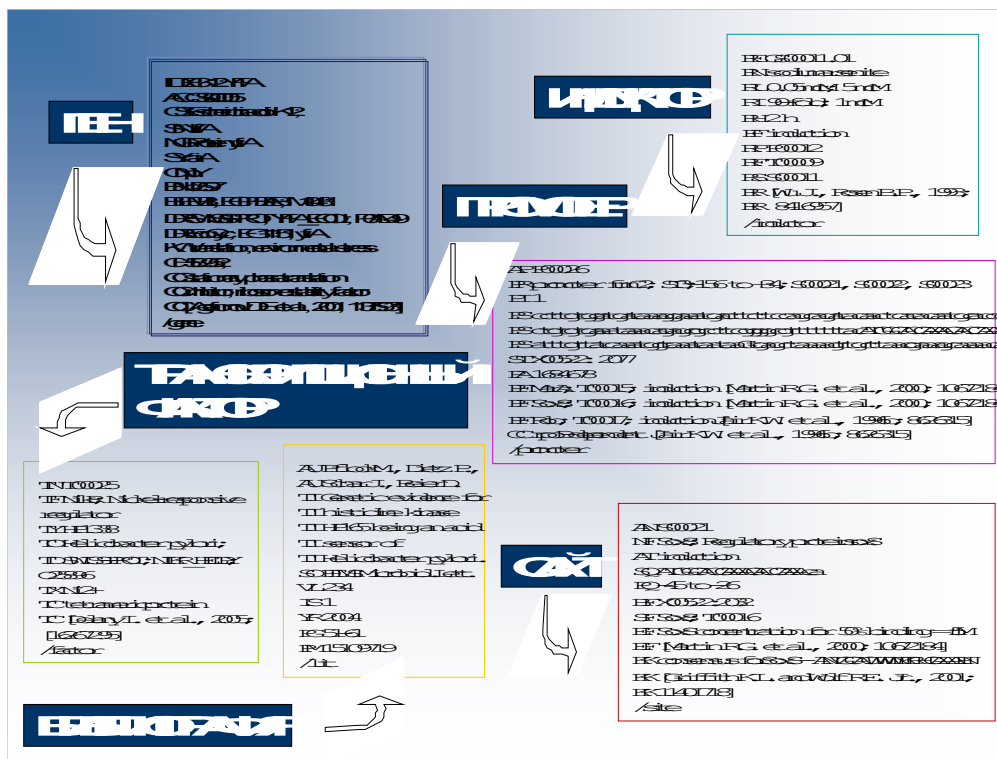


Рисунок 117. Информационные блоки базы GenSensor и форматы описания компонентов базы.

Структура входа в базу отражает организацию регуляторных районов бактериальных генов и содержит 61 различные поле, которые могут быть объединены в 6 информационных блоков: общее описание гена – 17 полей; описание промоторов – 9 полей; описание индукторов экспрессии гена – 11 полей; описание транскрипционных факторов – 8 полей; описание сайтов связывания транскрипционных факторов – 8. Кроме того, блок из 8 полей содержат информацию о публикациях, на основе которых создан данный вход.

Структура базы GenSensor и ее связи с внешними ресурсами отображены на рис. 118.

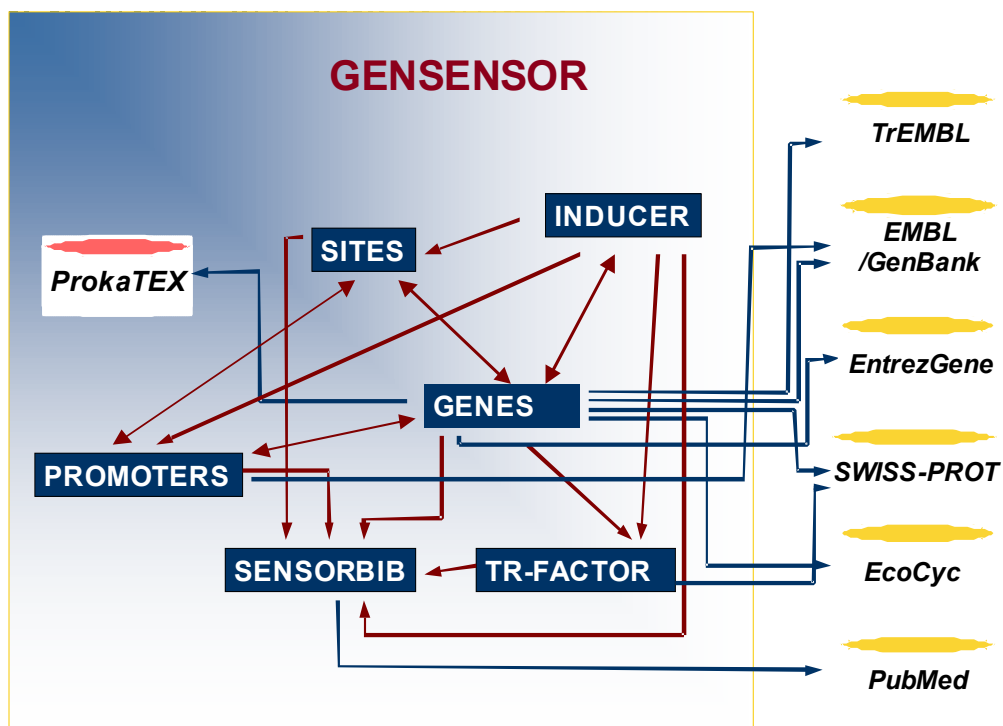


Рисунок 118. Структура базы GenSensor и ее связи с внешними ресурсами.

База может быть использована для поиска генов, тонко чувствующих присутствие в среде различных токсических для клетки веществ, и представляет интерес для исследователей, занимающихся разработкой и созданием новых, высокоэффективных и чувствительных методов, которые бы в отсутствие знаний о природе токсического для клетки вещества, выявляли его присутствие в различных средах.

В настоящее время база GenSensor содержит описание 128 промоторов генов *Escherichia coli*, *Helicobacter pylori* and *Bacillus subtilis*, экспрессия которых в существенной степени зависит от различного рода стрессовых воздействий, приводящих к повреждению ДНК, РНК, белков и мембран клеток. Влияние этих воздействий на экспрессию вышеприведенных генов отражено в 236 паттернах экспрессии.

Базовым программным средством, обеспечивающим поиск и навигацию по базе GenSensor, а также интеграцию GenSensor с внешними информационными и программными ресурсами, является Sequence Retrieval System (SRS). В соответствии с описанными выше информационными блоками, SRS представление базы GenSensor составляют 6 таблиц: GENSENSOR - SENSORGENE, SENSORINDUCTOR, SENSORPROMOTER, SENSORSITE, SENSORFACTOR, SENSORBIB. Система SRS обеспечивает возможность поиска необходимой информации по всем 6 таблицам. Технические подробности работы с базой данных GenSensor смотри «Руководство пользователя» АСНИ_01.

База данных ConSensor.

База данных ConSensor предназначена для накопления информации о существующих в научной литературе геносенсорных разработках и содержит данные о структуре и эффективности этих конструкций. Источником информации для базы является аннотация научных публикаций. Единицей входа в базу является эксперимент, в результате которого создана конкретная конструкция и показана ее функциональность. База ConSensor может

быть использована для поиска геносенсоров, чувствующих присутствие в среде различных токсических для клетки веществ и представляет интерес для исследователей, занимающихся разработкой и созданием новых, высокоэффективных и чувствительных методов мониторинга чистоты воды, воздуха, пищевых продуктов, а также контроля генотоксичности и/или мутагенности лекарственных препаратов, биологически активных добавок, бытовой химии и т.д. Эти знания дают не только возможность использования уже готовых разработок, но и возможность на их основе планировать более экономичным способом эксперименты по созданию новых конструкций, используя, в каком-то смысле, уже готовую методологию эксперимента.

База ConSensor только развивается и существует в виде флат-файла. Структура и формат базы ConSensor на примере описания геносенсорной конструкции на основе промотора гена *CadA* плазмиды *p1258 Staphylococcus aureus* и репортерного гена светлячка (*Photinus pyralis*) *lucFF* приведены на рисунке 119.

В настоящее время база содержит описание 76 генетических конструкций на основе промоторов генов *Escherichia coli*, *Staphylococcus aureus*, *Ralstonia metallidurans*, *Serratia marcescens*, and *Pseudomonas aeruginosa*. Их экспрессия резко возрастает в присутствии различных стрессовых факторов, вызывающих повреждение ДНК, РНК, белков и мембран. Влияние токсических для клетки веществ на уровень индукции биосенсоров, описанных в базе, отражено в 268 паттернах экспрессии.



Рисунок 119. Фрагменты описания в базе ConSensor геносенсорной конструкции на основе промотора гена *CadA* и репортерного гена *lucFF*, чувствительной к присутствию в среде тяжелых металлов.

Базовым программным средством, обеспечивающим поиск и навигацию по базе ConSensor, а также интеграцию ConSensor с внешними информационными и программными ресурсами, является Sequence Retrieval System (SRS). В соответствии с описанными выше информационными блоками, SRS представление базы ConSensor составляют три таблицы. Система SRS обеспечивает возможность поиска необходимой информации по всем трем

таблицам. Технические подробности работы с базой данных GenSensог смотри «Руководство пользователя» АСНИ_01.

3. СПИСОК ЛИТЕРАТУРЫ

1. [Демиденко Г.В., Колчанов Н.А., Лихошвай В.А., Матушкин Ю.Г., Фадеев С.И. // Математическое моделирование регулярных контуров генных сетей. Журнал вычислительной математики и математической физики, 2004, том 44, № 10, с. 1921–1940.](#)
2. [Колчанов Н.А., Ананько Е.А., Колпаков Ф.А., Подколodная О.А., Игнатъева Е.В., Горячкова Т.Н., Степаненко И.Л. Генные сети. Молекулярная биология, 2000, Т.34, № 4, с. 533-544.](#)
3. Корниш-Боуден Э. Основы ферментативной кинетики. М., Мир, 1979.
4. [Лихошвай В.А., Матушкин Ю.Г., Фадеев С.И. Задачи теории функционирования генных сетей. Журнал индустриальной математики, 2003, 6; 64-80.](#)
5. [Лихошвай В.А., Матушкин Ю.Г., Ратушный А.В., Ананько Е.А., Игнатъева Е.В., Подколodная О.В. Обобщенный химико-кинетический метод моделирования генных сетей. Молекулярная Биология 2001, т.35 \(6\), с. 1072-1079.](#)
6. [Лихошвай В.А., Фадеев С.И., Демиденко Г.В., Матушкин Ю.Г. Моделирование многостадийного синтеза вещества без ветвления уравнением с запаздывающим аргументом // Сибирский журнал индустриальной математики, 2004, том 7, №1, 73-94.](#)
7. Ратушный А.В., Лихошвай В.А., Игнатъева Е.В., Матушкин Ю.Г., Горянин И.И., Колчанов Н.А. Компьютерная модель генной сети регуляции биосинтеза холестерина в клетке: анализ влияния мутаций. Доклады Академии Наук, 2003, том 389, №2, с. 90-93.
8. [Ananko E.A., Podkolodny N.L., Stepanenko I.L., Ignatieva E.V. Podkolodnaya O.A., Kolchanov N.A. GeneNet: a database on structure and functional organization of gene networks. Nucleic Acids Research, 2002, V. 30, N 1, pp. 398-401.](#)
9. Ananko E.A., Podkolodny N.L., Stepanenko I.L., Podkolodnaya O.A., Rasskazov D.A., Miginsky D.S., Likhoshvai V.A., Ratushny A.V., Podkolodnaya N.N., Kolchanov N.A. (2005) GeneNet in 2005. *Nucleic Acids Res*, 33, D425-427.
10. Au N., Kuester-Schoeck E., Mandava V. et al. (2005) Genetic composition of the Bacillus subtilis SOS system. *J. Bacteriol.* 187, 7655-7666.
11. Brenda database: <http://brenda.bc.uni-koeln.de/>.
12. [Covert MW, Knight EM, Reed JL, Herrgard MJ, Palsson BO. Integrating high-throughput and computational data elucidates bacterial networks. Nature, 2004;429\(6987\):92-6.](#)
13. Dworkin M. "Prokaryotic Life Cycles," in M. Dworkin et al., eds., *The Prokaryotes: An Evolving Electronic Resource for the Microbiological Community*, 3rd edition. Springer-Verlag, 2001, New York.
14. [E. A. Ananko, N. L. Podkolodny, I. L. Stepanenko, O. A. Podkolodnaya, D. A. Rasskazov, D. S. Miginsky, V. A. Likhoshvai, A. V. Ratushny, N. N. Podkolodnaya and N. A. Kolchanov GeneNet in 2005. Nucleic Acids Res. 2005 Jan 1;33 \(Database issue\):D425-7.](#)
[Edwards R. Analysis of continuous-time switching networks // Physica D., 2000, vol. 146, pp. 165-199.](#)
Ellis L.B.M., Roe D., and Wackett L.P., "The University of Minnesota Biocatalysis/Biodegradation Database: The First Decade" *Nucleic Acids Res.*, 2006, 34, D517-D521.

15. [Fellenberg K., Hauser N.C., Brors B., Neutzner A., Hoheisel J.D., Vingron M. Correspondence analysis applied to microarray data. Proc Natl Acad Sci USA 2001; 98:10781-6.](#)
16. [Friedman N., Linial M., Nachman I., Pe'er D., 2000. Using Bayesian networks to analyze expression data // J. Comput. Biol., vol. 7, pp. 601-620.](#)
17. GenomeNet database: <http://www.genome.ad.jp/>.
18. Hakkila K., Maksimow M., Karp M., Virta M. (2002) Reporter genes lucFF, luxCDABE, gfp, and dsred have different characteristics in whole-cell bacterial sensors. *Anal. Biochem.* 301:235-242.
19. [Hofstadt R., Meineke F., 1995. Interactive modelling and simulation of biochemical networks // Comput. Biol. Med., vol. 25\(3\), pp. 321-334.](#)
20. [Kauffman S.A., 1993. The Origins of Order: Self-Organization and Selection in Evolution. N.Y.: Oxford Univ. Press.](#)
21. [N.A. Kolchanov, E.A. Ananko, V.A. Likhoshvai, O.A. Podkolodnaya, E.V. Ignatieva, A.V. Ratushny, Yu.G. Matushkin. Gene Networks Description and Modeling in the GeneNet System, Chapter 7, in "Gene Regulation and Metabolism", eds. Julio Collado-Vides and Ralf Hofstadt, The MIT Press, Cambridge, Massachusetts, 2002, 149-180.](#)
22. Keseler I.M., Collado-Vides J., Gama-Castro S., Ingraham J., Paley S., Paulsen I.T., Peralta-Gil M., Karp P.D. EcoCyc: a comprehensive database resource for Escherichia coli. *Nucleic Acids Res.* 2005, 33(Database issue):D334-337.
23. Khil P.P., Camerini-Otero R.D. (2002) Over 1000 genes are involved in the DNA damage response of Escherichia coli. *Mol. Microbiol.* 44(1):89-105.
24. [Ong I.M., Glasner J.D., Page D., 2002. Modelling regulatory pathways in E. coli from time series expression profiles // Bioinformatics, vol. 18, Suppl 1. P. 241-248.](#)
25. [Ratushny A.V., Ignatieva E.V., Likhoshvai V.A. Computer Dynamic Modeling of the Gene Network Controlling Intracellular Cholesterol Homeostasis In: Bioinformatics of genome regulation and structure. Ed. by N. Kolchanov and R. Hofstaedt, Kluwer Academic Publishers, Boston/Dordrecht/London, 2004, pp. 293-300.](#)
26. [Ratushny AV, Likhoshvai VA, Ignatieva EV, Kolchanov NA. Resilience of Cholesterol Concentration to a Wide Range of Mutations in the Cell. Complexus, 2003;1:142-148.](#)
Robison K., McGuire A.M., Church G.M. A comprehensive library of DNA-binding site matrices for 55 proteins applied to the complete Escherichia coli K-12 genome. *J. Mol. Biol.* 1998, 284(2):241-254.
27. Salgado H., Gama-Castro S., Peralta-Gil M., Diaz-Peredo E., Sanchez-Solano F., Santos-Zavaleta A., Martinez-Flores I., Jimenez-Jacinto V., Bonavides-Martinez C., Segura-Salazar J., Martinez-Antonio A., Collado-Vides J. RegulonDB (version 5.0): Escherichia coli K-12 transcriptional regulatory network, operon organization, and growth conditions. *Nucleic Acids Res.* 2006, 34(Database issue):D394-397.
28. Schweigert S., Herde P.V., Sibbald P.R. (1995) Issues in incorporation semantic integrity in molecular biological object-oriented databases. *Comput Appl Biosci*, 11, 339-347.
29. [Sherlock G, Hernandez-Boussard T, Kasarskis A, Binkley G, Matese JC, Dwight SS, Kaloper M, Weng S, Jin H, Ball CA, Eisen MB, Spellman PT, Brown PO, Botstein D, Cherry JM.. The Stanford microarray database. Nucleic Acids Res 2001; 29:152-155.](#)
30. [Smolen P., Baxter D.A., Byrne J.H., 2000. Modeling transcriptional control in gene networks - methods, recent results, and future directions // Bull. Math. Biol., vol. 62\(2\), pp. 247-292.](#)
31. Stephanopoulos G. Metabolic engineering by genome shuffling. *Nat. Biotechnol.*, 2002, 20, 666-668.

32. [Tchuraev R.N., 1991. A new method for the analysis of the dynamics of the molecular genetic control systems. I. Description of the method of generalized threshold models // J. Theor. Biol., vol. 151\(1\), pp. 71-87.](#)
33. [Thomas R., Thieffry D., Kaufman M., 1995. Dynamical behaviour of biological regulatory networks--I. Biological role of feedback loops and practical use of the concept of the loop-characteristic state // Bull. Math. Biol., vol. 57\(2\), pp. 247-276.](#)
34. [Turner T.E., Schnell S., Burrage K., 2004. Stochastic approaches for modelling in vivo reactions // Comput. Biol. Chem., vol. 28\(3\), pp. 165-178.](#)
35. Urbance J.W., Cole J., Saxman P., Tiedje J.M. BSD: the Biodegradative Strain Database, *Nucleic Acids Res.*, 2003, 31, 152-155.
36. Zheng M., Wang X., Templeton L.J., Smulski D.R., LaRossa R.A., Storz G. (2001) DNA microarray-mediated transcriptional profiling of the *Escherichia coli* response to hydrogen peroxide. *J. Bacteriol.* 183(15):4562-4570.