

(19)



Евразийское
патентное
ведомство

(11) 035276

(13) B1

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ

(45) Дата публикации и выдачи патента
2020.05.22

(21) Номер заявки
201791659

(22) Дата подачи заявки
2016.01.28

(51) Int. Cl. G06F 19/12 (2011.01)
G06F 19/26 (2011.01)

(54) КОМПЬЮТЕРИЗОВАННЫЙ СПОСОБ СОЗДАНИЯ МОДЕЛИ ПРОЦЕССА
ФЕРМЕНТАЦИИ

(31) 15153052.4

(32) 2015.01.29

(33) EP

(43) 2018.01.31

(86) PCT/EP2016/051753

(87) WO 2016/120361 2016.08.04

(71)(73) Заявитель и патентовладелец:
БАЙЕР АКЦИЕНГЕЗЕЛЬШАФТ
(DE)

(72) Изобретатель:
Нейманн Тобиас, Хебинг Лукас,
Энгель Себастиан (DE)

(74) Представитель:
Беляева Е.Н. (BY)

(56) J. GAO ET AL.: "Dynamic Metabolic Modeling for a MAB Bioprocess", BIOTECHNOLOGY PROGRESS, vol. 23, no. 1, 2 February 2007 (2007-02-02), pages 168-181, XP055199374, ISSN: 8756-7938, DOI: 10.1021/bp060089y abstract, page 169, column 1, paragraph 3 - page 170, column 1, paragraph 1, page 175, column 1, paragraph 1 - paragraph 3

PROVOST A. ET AL.: "Metabolic design of macroscopic bioreaction models: application to Chinese hamster ovary cells", BIOPROCESS AND BIOSYSTEMS ENGINEERING, SPRINGER, BERLIN, DE, vol. 29, no. 5-6, 30 September 2006 (2006-09-30), pages 349-366, XP019461015, ISSN: 1615-7605, DOI: 10.1007/S00449-006-0083-Y abstract, page 350, column 2, paragraph 1 - page 352, column 2, paragraph 5, figures 1,6, page 171, column 2, paragraph 3 - page 173, column 1, paragraph 1

(57) Изобретение касается компьютеризованного способа создания модели биореакции, в частности процесса ферментативного или целлюлозного катализа, с использованием организма на основе данных измерений.

035276 B1

035276 B1

035276 B1

Изобретение касается компьютеризованного способа создания модели биореакции, в частности ферментативного или целлюлозного катализа, с использованием организма.

Организмом в значении данной заявки являются культуры из клеток растений или животных, такие как клетки млекопитающих, дрожжи, бактерии, водоросли и т.д., которые используют в биореакциях.

Сенсорный контроль процесса ферментации и анализ образцов из процесса, например с помощью аналитической автоматизированной платформы BaychromAT® Quality-by-Design (концепт качества на этапе разработки) компании Bayer Technology Services GmbH показывает различные данные о состоянии процесса в биореакторе в режиме реального времени. Обычно устанавливают количество клеток, жизнеспособность клеток, концентрацию субстрата как источника углерода (например, глюкоза), аминокислоты или O_2 , вещества и побочные продукты (например, лактат или CO_2), параметры процесса как температура, и/или значение pH, или признаки веществ. Эти данные могут быть дополнены расчетными данными и/или экстраполяциями, например согласно уровню техники. Вместе эти данные образуют данные измерений или технологические данные в значении данной заявки.

Базовые сведения об организме в значении данной заявки означают сведения о биохимических реакциях организма (специфических и неспецифических реакциях) и особенно внутриклеточных реакциях, или макрореакциях для описания особых для организма сетей обмена веществ (также называемых SN или метаболических сетей), которые состоят из субстратов, метаболитов (называемых также узлами метаболической сети), продуктов, а также биохимических реакций между ними. Эти биохимические реакции определяют их

- (a) стехиометрия,
- (b) обратимость реакций (в биологических условиях),
- (c) интеграция в стехиометрическую сеть.

До сих пор данные измерений применяют в основном для контроля качества процесса. Далее представлен выбор технических постановок вопроса, для решения которых требуются динамические модели процесса.

Техническое применение технологических данных в значении данной заявки предлагает оценку процесса в реакторе, основанную на моделях. Такие методы как метод с использованием фильтра Калмана (англ. extended Kalman Filter), способствуют непрерывной оценке масштаба процесса с помощью периодических измерений [Welch G, Bishop G. 1995. An Introduction to the Kalman Filter. Chapel Hill, NC, США: University of North Carolina at Chapel Hill]. Также можно рассчитывать диаграмму неизмеримых величин на основании других измерений. Условием для этого является правильно описанная модель процесса, лежащая в основе процесса.

Другим применением является проведение основанного на модели оптимального технологического процесса. При этом применяют динамическую модель процесса, чтобы в основанном на модели, предиктивном, закрытом контуре регулирования оптимизировать технологический процесс относительно количества и признаков продукта, или образования побочных продуктов, или других целевых функций. Это, например, показывает Frahm и др. для культуры клеток гибридома [Frahm B, Lane P, Atzert H, Munkack A, Hoffmann M, Hass VC, Portner R. 2002. Adaptive, Model-Based Control by the Open-Loop-Feedback-Optimal (OLFO) Controller for the Effective Fed-Batch Cultivation of Hybridoma Cells. Biotechnol. Prog. 18(5):1095-1103].

Для обоих названных технических применений важным является то, что получаемая модель процесса имеет минимальную степень сложности, т.е. ограниченное количество фазовых переменных и/или уравнений, при одновременно хорошей точности воспроизведения процесса. Кроме названных применений для технологического процесса можно также использовать динамические модели процесса во время его разработки, чтобы планировать эксперименты с оптимальным получением информации. Эти действия называют "основанное на модели планирование эксперимента" [Franceschini G, Macchietto S. 2008. Model-based design of experiments for parameter precision: State of the art. Chemical Engineering Science 63(19):4846-4872]. Кроме вышеназванных условий для комплексности модели они необходимы для этого технического применения, динамическая модель процесса уже присутствует во время фазы развития. Ее необходимо получить как можно быстрее из уже присутствующих технологических данных, чтобы затраты времени на развитие процесса были минимальными.

Поэтому возникла необходимость в получении способа, который способствует образованию динамической модели процесса с использованием базовых сведений и технологических данных. Чтобы можно было использовать эту модель, например, для оценки состояния, оптимального технологического процесса или основанного на модели планирования эксперимента, необходимо, чтобы степень сложности модели была минимальной. Зависимость, т.е. влияние величины процесса или состояния процесса на его поведение необходимо достаточно точно квантифицировать внутри пространства проектных решений. Для этого необходимо использовать всю имеющуюся информацию о состоянии процесса. Основанное на модели описание признаков продукта при необходимости должно быть интегрируемым в модель. В качестве пространства проектных решений присутствует область, названная "технологические данные". Способ должен быть применим в вышеназванных биореакциях, и время, необходимое для развития такой динамической модели, должно быть значительно уменьшено. Предыдущие разработки для разви-

тия динамических моделей требуют от нескольких месяцев или лет до получения модели процесса. Данная разработка согласно изобретению сокращает период разработки до нескольких недель.

Типичными признаками продукта в значении данной заявки являются, например, образец гликозилирования протеинов или целостность протеинов, но не ограничиваются ими. Динамические модели, которые используют в вышеназванном контексте, до сих пор не обнаруживают этих свойств. Данная разработка способствует простой, основанной на модели интеграции признаков продукта.

Основанный на модели технологический процесс ферментаций показали Frahm и др. на примере культуры клеток гибридомы (Frahm B, Lane P, Atzert H, Munack A, Hoffmann M, Hass VC, Portner R. 2002. Adaptive, Model-Based Control by the Open-Loop-Feedback-Optimal (OLFO) Controller for the Effective Fed-Batch Cultivation of Hybridoma Cells. *Biotechnol. Prog.* 18(5): 1095-1103). Основными параметрами процесса управляют здесь на основе модели. Интеграция признаков продукта здесь не происходит. Для этого специфического процесса спроектировали математическую модель клетки, которую можно переводить с использованием больших затрат на процессы с такими же или другими организмами или штаммами одинакового организма. Базовые сведения в виде внутриклеточных реакций в модели не учитывают. Интеграцию других показателей в модель и, таким образом, полное использование информации о состоянии процесса можно осуществлять только с использованием больших затрат. Таким образом, разработка представляет собой индивидуальное решение, которое не переносится на другие процессы и не способствует полному использованию определенных данных. Данный метод из-за ожидаемого периода, необходимого для разработки модели, и дорогостоящей переносимости решения на другие процессы с такими же или другими организмами не решает вышеназванной технической проблемы.

Дальнейшее моделирование, которое также включает в себя такие признаки продукта как гликозилирование, можно найти в работах Kontoravdi и др. Модель, которая описывает основной обмен веществ, не включает в себя базовых сведений в виде внутриклеточных реакций и не может быть перенесена на другие процессы с такими же или другими организмами. Здесь нельзя проводить интеграцию других показателей в модель [Kontoravdi C, Asprey SP, Pistikopoulos EN, Mantalaris A. 2007. Development of a dynamic model of monoclonal antibody production and glycosylation for product quality monitoring. *Computers & Chemical Engineering* 31(5-6):392-400]. Этот метод также не способствует полному использованию информации о состоянии процесса, требует большого количества времени для разработки модели и не может быть перенесен на другие организмы или штаммы. Таким образом, этот метод не является решением технической проблемы.

Моделирование гликозилирования с интеграцией обмена нуклеотидных сахарных веществ, описанное Jedrzejewski и др., и Jimenez, и др., включает в себя базовые сведения в виде уравнения баланса внутренних промежуточных продуктов обмена веществ [Jedrzejewski PM, del Val, Ioscani Jimenez, Constantinou A, Dell A, Haslam SM, Polizzi KM, Kontoravdi C. 2014. Towards Controlling the Glycoform: A Model Framework Linking Extracellular Metabolites to Antibody Glycosylation. *International journal of molecular sciences* 15(3):4492-4522.; Jimenez del Val, Ioscani, Nagy JM, Kontoravdi C. 2011. A dynamic mathematical model for monoclonal antibody N-linked glycosylation and nucleotide sugar donor transport within a maturing Golgi apparatus. *Biotechnology progress* 27(6): 1730-1743]. При использовании этой модели для технологического процесса недостатками являются сложность всей модели и недостаточная наблюдаемость за внутренними промежуточными продуктами обмена веществ. Кроме того, модель основного обмена веществ нельзя переносить на другие процессы или полностью использовать информацию о состоянии технического процесса. Таким образом, этот метод не является решением технической проблемы.

Гибкую генерацию модели биопроцесса предлагают Leifheit и др. [Leifheit J, Heine T, Kawohl M, King R. 2007. Компьютерное полуавтоматическое моделирование биотехнологических процессов (Semi-automatic Modeling of Biotechnical Processes) - *Техника автоматизации* 55(5)]. Генерацию модели проводят с помощью технологических знаний, но не используя базовые сведения. Процедура применима для различных процессов с одинаковыми или разными организмами. Ее основой являются макрореакции, которые заданы самим пользователем. Их точную стехиометрию определяют во время способа. Этот метод описан для незначительного количества параметров или показателей. Интеграция других параметров или показателей была бы связана со значительным увеличением сложности способа. При использовании значительной основы данных, которые предоставляет, например, BaychoMAT®-платформа, этот метод был бы невыполним. Метод не позволяет интеграцию признаков продукта. Таким образом, он не является решением вышеназванной технической проблемы.

Применение базовых сведений в виде макрореакций, которые были получены в виде элементарных мод (англ.: Elementary Modes) (EM) из известных метаболических (стехиометрических) сетей организма, описал Provost [Provost A. 2006. Metabolic design of dynamic bioreaction models. *Faculté des Sciences Appliquées, Université catholique de Louvain, Louvain-la-Neuve, Louvain-la-Neuve*, стр. 81 и др., стр. 107 и др., стр. 118 и др.]. Этот метод применим для различных организмов или штаммов одинаковых организмов. Выбор макрореакций для модели процесса осуществляют с применением технологических знаний. Однако отдельные участки процесса, определенные для заданного количества макрореакций, выбраны случайно. Описанный способ позволяет использование множества различных комбинаций элементарных мод (EM). Количество макрореакций и, таким образом, сложность модели являются определенными и не

могут меняться. Способ показывает отдельные модели для каждого участка процесса. Выбор кинетики для отдельных макрореакций проводят, принимая во внимание стехиометрию выбранных макрореакций. Однако параметры кинетики (параметры модели) не согласованы с технологическими данными. Вместо этого применение отдельных моделей участка процесса отражает изменения в технологических данных. Хотя случайный выбор реакций можно также проводить на основе обширных базовых данных, с помощью описанной разработки для выбора кинетик и выбранных кинетик нельзя отобразить ход процесса или поведение организма в процессе. Кроме того, применение нескольких моделей участка процесса приводит к ненужному повышению сложности модели процесса. В этом методе не дают количественную оценку зависимости, т.е. воздействию величины процесса или состояния процесса на его развитие. Также здесь не происходит интеграция признаков продукта. Таким образом, этот метод не является решением вышеописанной технической проблемы.

Поэтому возникает необходимость в получении способа, который способствует быстрому и эффективному получению модели на основе технологических знаний и данных и оптимизации превращений продукта и критических признаков продукта, принимая во внимание базовые сведения, и который не обнаруживает вышеуказанных недостатков.

Задачу решили с помощью способе получения модели биореакции с одним организмом в одном биореакторе, как описано ниже.

Предметом заявки является компьютеризованный способ получения модели биореакции, в частности ферментативного или целлюлозного катализа, с использованием организма, который включает следующие этапы.

а) Выбранные пути обмена веществ организма, стехиометрические свойства, а также свойства обратимости реакций, которые вводят в способ в виде базовых сведений. Говоря другими словами, в способ вводят одну или более метаболических сетей организма. Элементарные моды (ЕМ) рассчитывают из этих данных.

б) ЕМ объединяют в матрице К, причем ЕМ объединяют пути обмена веществ из этапа а) в макро-реакциях. Вместе с этим указанная матрица К содержит стехиометрические свойства, а также свойства обратимости реакций всех возможных макрореакций из базовых сведений.

с) Данные измерений (называемые также "технологическими данными") вводят для биореакции с использованием организма.

д) С помощью метода интерполяции на основе предоставленных данных измерений из этапа с) рассчитывают специальные для организма скорости реакций (скорость выхода и скорость поступления одного или более входных параметров и выходных параметров) для введенных путей обмена веществ. Также предпочтительно рассчитывают скорость роста, особенно предпочтительно также показатели гибели микроорганизмов.

е) Соответствующие макрореакции выбирают в виде подмножества элементарных мод из этапа б) посредством

i) независимого и/или зависимого от данных предварительного сокращения количества ЕМ из этапа б);

ii) выбора подмножества из предварительного сокращения из этапа е)-i) с данными измерений из этапа с) и/или одной или более скоростей реакций из этапа д), предпочтительно с данными измерений из этапа с) с помощью алгоритма согласно математическому критерию качества и объединению подмножества в матрице L;

iii) при необходимости подмножества отображают графически.

ф) С помощью метода интерполяции на основе предоставленных данных измерений из этапа с) и/или скоростей реакций из этапа д) рассчитывают скорости реакций макро-реакций подмножества $L(t)$.

г) Кинетики макро-реакций подмножества е)-ii) проектируют на следующих промежуточных этапах, благодаря этому определяют параметры модели:

i) из стехиометрии макро-реакций проектируют родовые кинетики;

ii) параметры воздействия на макро-реакции определяют из скоростей реакций из этапа ф);

iii) родовые кинетики из этапа г)-i) пополняются элементами, которые дают количественную оценку установленным на этапе г)-ii) входным параметрам.

h) При необходимости, для кинетик из этапа г) проводят первое согласование значений параметров модели отдельно для каждой макро-реакции с рассчитанными скоростями реакций из этапа ф) и проверяют это согласование.

i) При необходимости, повторяют этапы г) и h) до получения заданного согласования.

j) Значения параметров модели согласуют с данными измерений из этапа с).

к) Матрица L, кинетики из этапа г) и значения параметров модели из этапа j) создают модель и вводят и/или передают в модуль управления технологическим процессом или разработки технологического процесса.

Как правило, модуль управления технологического процесса сообщается онлайн с системой управления технологическим процессом, которую обычно используют для управления биореактором.

Как правило, модуль разработки технологического процесса применяют для офлайн-оптимизации

процесса или планирования экспериментов.

Моделирование согласно изобретению биореакции основывается в основном на принятии репрезентативных макрореакций, которые представляют собой внутренние процессы обмена веществ. Для выбора реакций необходимы как биохимические базовые сведения, так и технические знания.

На первом этапе способа проводят введение реакций метаболической сети, ее стехиометрии и свойств обратимости реакций пользователем через рабочую поверхность или в идеальном случае автоматически с помощью выбора организма и заложенных в нем путей обмена веществ из модуля банка данных, в котором собраны базовые сведения об организме. Метаболическая сеть (в уровне техники также называемая "стехиометрическая сеть") и свойства ее отдельных реакций представляют собой базовые сведения организма. Предпочтительно метаболическая сеть содержит реакции из важных для организма путей обмена веществ, например реакций гликолиза. Особенно предпочтительно выбор содержит внешние реакции. Внешняя реакция в значении данной заявки содержит по меньшей мере один компонент, кроме клетки, обычно по меньшей мере один параметр и/или по меньшей мере один выходной параметр (продукт, побочный продукт и т.д.). Особенно предпочтительно метаболическая сеть содержит реакции, которые описывают рост клеток, например в виде упрощенной реакции внутренних метаболитов к внешней биомассе. Фиг. 5 и табл. 1 в примере описывают применимую метаболическую сеть, не ограничиваясь этим.

Затем из введенных, объединенных в одну или более стехиометрических сетей путей обмена веществ организма рассчитывают элементарные моды. Каждая элементарная мода является линейной комбинацией скоростей реакций из путей обмена веществ, т.е. внутренних и внешних реакций метаболической сети, в которой как выполняется условие установившегося режима (англ. steady state) для внутренних метаболитов, так и учитывается обратимость или необратимость реакций. При линейной комбинации реакций, которые учитывают условие установившегося режима для внутренних метаболитов, не могут накапливаться никакие внутренние метаболиты.

Внутренняя реакция в значении заявки проходит только внутри клетки.

Через экстернализацию внутреннего компонента, т.е. с помощью классификации собственно внутреннего компонента в качестве входного или выходного параметра, можно моделировать связанную с экстернализованными внутренними компонентами реакцию в виде внешней реакции и таким образом обойти условие установившегося режима для внутренних метаболитов в этом случае.

Макрореакция в значении заявки содержит все реакции, которые от одного или более входных параметров приводят к одному или нескольким выходным параметрам. Каждая элементарная мода описывает таким образом макрореакцию. По сравнению с методом Leifheit и др. в значении изобретения получают макрореакции на основе введенных базовых сведений.

Элементарная мода (ЕМ) интегрирована в матрицу E, предпочтительно в модуле для конструкции матрицы, который расположен с соответствующим алгоритмом. Для постройки матрицы элементарных мод можно использовать известные алгоритмы. METATOOL назван в качестве примера, но им не ограничиваются: [Pfeiffer T, Montero F, Schuster, стр. 1999. METATOOL: for studying metabolic networks. Bioinformatics 15(3):251-257.]

С помощью METATOOL получают первую матрицу E, которая описывает введенные внутренние и внешние реакции.

На этапе b) с помощью (внешней) стехиометрической матрицы N_p из матрицы (E) получают матрицу, состоящую из возможных макрореакций K.

$$K = N_p \cdot E \text{ (формула 1)}$$

Трансформация матрицы E после K известна согласно Provost [Provost A. 2006. Metabolic design of dynamic bioreaction models. Faculté des Sciences Appliquées, Université catholique de Louvain, Louvain-la-Neuve. Louvain-la-Neuve, стр. 81].

Векторы-столбцы матрицы описывают макрореакции. Векторы-строки описывают компоненты макрореакций (входные и выходные параметры). В матрицу K вносят стехиометрию макрореакций.

Каждая возможная в значении метаболической сети скорость реакции может быть представлена в виде положительной линейной комбинации этих макрореакций.

Применение ЕМ в виде основы модели процесса известно в уровне техники (например, из Provost A. 2006. Metabolic design of dynamic bioreaction models. Faculté des Sciences Appliquées, Université catholique de Louvain, Louvain-la-Neuve. Louvain-la-Neuve, стр. 87, стр. 118 и др.; Gao, J. и др. (2007). Dynamic metabolic modeling for a MAB bioprocess. Biotechnology progress, 23(1), 168-181).

На следующем этапе c) вводят имеющиеся в наличии данные измерений (технические данные) для биореакции с использованием организма. Обычно устанавливают количество клеток, жизнеспособность клеток, концентрацию субстрата как источника углерода (например, глюкоза), аминокислоты или O_2 , вещества и побочные продукты (например, лактат или CO_2), параметры процесса как температура, и/или значение pH, или признаки веществ. Это введение можно производить пользователем вручную или автоматически, как, например, с помощью селекции из модуля банка данных для хранения данных измерений и перевода выбранных данных в модуль анализа данных, который связан с модулем банка данных.

Из этих данных измерений на этапе d) рассчитали специфичные для клеток скорости выхода и по-

ступления субстратов и продуктов (побочных продуктов) - названные вместе специфические скорости $q(t)$, а также при необходимости скорости роста и показателей гибели микроорганизмов организма ($\mu(t)$, $\mu_d(t)$).

Условием для расчета является интерполяция количества живых клеток, общего числа клеток и концентрации среды с помощью метода интерполяции. Из этого можно установить временные изменения измеряемых параметров. Рассчитанные нормы $q(t)$, $\mu(t)$, $\mu_d(t)$ дают сведения о наблюдаемом динамическом поведении организма в течение времени.

Для расчета вышеуказанных данных можно использовать один или несколько различных методов интерполяции в комбинации.

Например, Leifheit и др. описывают вычисление временных изменений измеряемых параметров, например вычисление общего числа клеток, вычисление количества живых клеток или вычисление других концентраций среды, из данных измерений с помощью интерполированных сплайнами данных измерений [Leifheit, J., Heine, T., Kawohl, M., & King, R. (2007). Компьютеризованное полуавтоматическое моделирование биотехнологических процессов (Semiautomatic Modeling of Biotechnical Processes). *at-Automatisierungstechnik*, 55(5), 211-218]. Вместе с этим этот метод включен в данную заявку посредством ссылок.

Вышеуказанные нормы $q(t)$, $\mu(t)$, $\mu_d(t)$ рассчитывают из этих временных изменений:

Например, скорость роста организма $\mu(t)$ можно рассчитать из интерполированных сплайнами значений общего числа клеток $X_t(t)$ и живых клеток $X_v(t)$ а также из рассчитанного из них временного изменения общего числа клеток $\frac{dX_t}{dt}(t)$ с формулой 2:

$$\frac{dX_t}{dt}(t) = -D(t) \cdot X_t(t) + \mu(t) \cdot X_v(t) \quad (\text{формула 2})$$

Причем $D(t)$ означает скорость разбавления.

Показатель гибели микроорганизмов $\mu_d(t)$ можно рассчитать при известном направлении $\mu(t)$ из направления $X_v(t)$ и направления временного изменения количества живых клеток $\frac{dX_v}{dt}(t)$ с формулой 3:

$$\frac{dX_v}{dt}(t) = -D(t) \cdot X_v(t) + (\mu(t) - \mu_d(t)) \cdot X_v(t) \quad (\text{формула 3})$$

Расчет специфических норм других компонентов i $q_i(t)$ можно получить из интерполированных сплайнами значений количества живых клеток $X_v(t)$ и концентрации компонента $C_i(t)$, а также из направления временного изменения $\frac{dC_i}{dt}(t)$, из интерполированных сплайнами значений $C_i(t)$ с помощью формулы 4:

$$\frac{dC_i}{dt}(t) = D(t) \cdot (C_{i, in} - C_i(t)) + q_i(t) \cdot X_v(t) \quad (\text{формула 4})$$

В предпочтительной форме выполнения способа предварительно обрабатывают данные измерений из этапа с) перед первой интерполяцией следующим образом: Чтобы учитывать все, вызванные не клетками изменения концентрации и из данных измерений получить постоянное направление изменений концентрации, данные измерений смещают (в данной заявке называется сдвиг). Сумму $\Delta C_i(t)$, на которую смещается изменение концентрации, можно рассчитать по формуле 5:

$$\Delta C_i(t) = \int_0^t D(\tau) \cdot (C_{i, in} - C_{i, s}(\tau)) d\tau \quad (\text{формула 5})$$

Причем $D(\tau)$ означает скорость разбавления. Направление концентрации со сдвигом $C_{i, s}(t)$ получается по формуле 6:

$$C_{i, s}(t) = C_i(t) - \Delta C_i(t) \quad (\text{формула 6})$$

Дифференциальное уравнение, которое показывает направление концентрации со сдвигом $C_{i, s}(t)$, получается из формул 4 и 6:

$$\frac{dC_{i, s}}{dt} = \frac{d(C_i(t) - \Delta C_i(t))}{dt} = q_i(t) \cdot X_v(t) \quad (\text{формула 7})$$

Это редактирование (сдвиг) данных измерения предотвращает резкое изменение рассчитанных специфических норм при подаче и выходе подаваемого материала (пик подаваемого материала), особенно в периодическом процессе с подпиткой.

Фиг. 1 показывает предварительную обработку/сдвиг данных измерений в значении данной заявки.

В особенно предпочтительной форме выполнения способа используют предварительно обработанные данные для вычисления градиентов общего числа клеток $\frac{dX_t}{dt}(t)$ с помощью метода Leifheit и др.

Это сходится с интерполяцией сплайнами согласно дифференциальному уравнению 8:

$$\frac{dX_t}{dt}(t) = \mu(t) \cdot X_v(t) \quad (\text{формула 8})$$

Особенно предпочтительно включают лизис в расчеты на основании фактора лизиса K_l (формула 9).

Это можно принимать, например, постоянно в течение технологического процесса.

$$\left. \frac{dX_t}{dt} \right|_s = \mu(t) \cdot X_v(t) - K_l \cdot (X_t(t) - X_v(t)) \quad (\text{формула 9})$$

Уменьшение общего числа клеток со сдвигом $X_{t,s}(t)$ можно объяснить с помощью лизиса таким образом, что можно предотвратить отрицательные значения для скорости роста $\mu(t)$.

Предпочтительно предварительно обработанные данные также используют для расчета показателя гибели микроорганизмов $\mu_d(t)$.

Предпочтительно представлены возможные комбинации специфических скоростей $q(t)$ в диаграмме Flux-Мар, см., например, фиг. 2. Это изображение дает обзор рассчитанных специфических скоростей $q(t)$. Горизонтالي показывают, какие области являются физиологически важными.

Если специфические скорости $q(t)$ и при необходимости другие показатели $\mu(t)$ и $\mu_d(t)$ имеют различные порядки величин и единицы, то их обычно объединяют с помощью упрощений в специфический вектор норм $\tilde{q}(t)$ с одинаковыми единицами. Например, специфическая норма макромолекулы, которая

измеряется в граммах [g], имеет единицу $\left[\frac{g}{\text{Cell} \cdot h} \right]$. Если оценивают состав этой макромолекулы, основанный, например, на ее С-моль-содержании, то можно изменить ее специфическую норму с [g] на [C-mol], таким образом специфическая норма имеет единицу $\left[\frac{\text{C-mol}}{\text{Cell} \cdot h} \right]$.

Специфические нормы $\tilde{q}(t)$ образуют одну из основ для следующего этапа е) способа, а именно выбор важных макрореакций.

На этапе е) выбирают подмножество (L) ЕМ на основе данных, с помощью которого можно хорошо отобразить специфические нормы $\tilde{q}(t)$ из этапа d) и/или данные измерений из этапа с) согласно одному математическому критерию качества. Количество ЕМ в подмножестве (L) должно быть максимально низким, чтобы обеспечить наиболее низкую сложность модели процесса. Подмножество L также должно обеспечить хорошее описание технологических знаний.

Избирательность ЕМ уменьшает пространство решений по сравнению с оригинальным ЕМ-комплексом (K) из этапа а), но также содержит установленную физиологически важную область клеток.

Фиг. 3 показывает изображение пространства решений, причем снижают оригинальный комплект ЕМ (K) к подмножеству (L).

Для этапа е) обычно передают рассчитанные специфические нормы $\tilde{q}(t)$, а также данные измерений из этапа с) в модуль для выбора важных макрореакций, который создает конфигурацию с соответствующим алгоритмом.

На этапе е) i) проводится независимая от данных и/или зависимая от данных предварительная редукция матрицы K в любом порядке:

Независимую от данных предварительную редукцию обычно проводят с помощью геометрической редукции. При этом для случайно выбранных ЕМ рассчитывают все косинус-подобия ко всем другим режимам. ЕМ с наибольшим подобием t удаляют из комплекта. Этот способ действия повторяют, пока не получают предварительно определенное количество ЕМ. Желаемое количество обычно определяют для способа в предзоне. В качестве контрольной переменной можно использовать объем пространства решений. Неожиданно было установлено, что возможно значительное сокращение количества макрореакций, сохраняя 90-98%, предпочтительно 92-95% закрепленного объема по сравнению с оригинальным объемом.

Зависимую от данных предварительную редукцию можно проводить с помощью сравнения коэффициентов выхода продукции ЕМ (Y^{EM}) к коэффициентам выхода продукции, которые рассчитывают из специфических норм $\tilde{q}(t)$ из этапа d) (Y^m). Коэффициент выхода продукции k-t-ЕМ ($Y_{i,j}^{EM,k}$) устанавливают по формуле 10 с помощью деления соответствующих стехиометрических коэффициентов внешних метаболитов i и j. Для k-t ЕМ это внесение матрицы $K_{i,k}$ и $K_{j,k}$.

$$Y_{i,j}^{EM,k} = \frac{K_{i,k}}{K_{j,k}} \quad (\text{формула 10})$$

Если стехиометрический коэффициент $K_{j,k}=0$, то коэффициент выхода продукции нельзя установить.

Коэффициент выхода продукции $Y_{i,j}^m(t)$ дает соотношение дважды определенное или рассчитанное согласно d) из особых клеточных норм $(\tilde{q}_i(t), \tilde{q}_j(t))$ друг к другу по формуле 11:

$$Y_{i,j}^m(t) = \frac{\tilde{q}_i(t)}{\tilde{q}_j(t)} \quad (\text{формула 11})$$

Из коэффициентов выхода продукции Y^m для каждой возможной комбинации двух внешних компонентов i и j можно определить внешнюю и нижнюю границу. Например, можно использовать самый

маленький коэффициент выхода продукции двух внешних метаболитов i и j $Y_{i,j}^m(t)$ в качестве нижней границы и самое большое значение $Y_{i,j}^m(t)$ в качестве верхней границы, однако возможно использование и других границ. ЕМ, коэффициенты выхода продукции $Y_{i,j}^{EM}$ которых находятся выше верхней границы или ниже нижней границы, удаляют из матрицы K . Если коэффициент выхода продукции ЕМ $Y_{i,j}^{EM}$ нельзя определить, то он остается в матрице K . Предпочтительно также можно использовать описанный на странице 15 метод согласно изобретению "линейная оценка норм реакции выбранных макрореакций с NNLS" для зависимой от данных предварительного сокращения. Там находится пояснение метода в контексте с предварительной редукцией. Преимуществом вовлечения технологических данных в зависимой от данных предварительной редукции в способе является то, что редукция приведена в процесс и таким образом осуществляется более эффективно и фокусированно.

На этапе е) ii) выбирают подмножество из макрореакций с алгоритмом. Для выбора требуется критерий качества, с помощью которого можно дать количественную оценку того, насколько хорошо можно отобразить специфические нормы $\tilde{q}(t)$ из этапа d) и/или данные измерений из этапа с) с подмножеством (L) , а также алгоритм для выбора подмножества.

В качестве критерия качества для изображения рассчитанных специфических норм $\tilde{q}(t)$ с подмножеством L можно использовать согласно Soons и др. сумму возведенного в квадрат остатка специфических норм (SSR_q) по формуле 12 [Soons, Z.I.T.A., Ferreira, E.C., Rocha, I. (2010). Selection of Elementary Modes for Bioprocess Control. 11th International Symposium on Computer Applications in Biotechnology, Лёвен, Бельгия, июль 7-9, 2010, 156-161].

$$SSR_q = \sum_{i=1}^{N_t} \|L \cdot \underline{r}(t_i) - \tilde{q}(t_i)\| \quad (\text{формула 12})$$

Значение для SSR_q должно быть максимально низким.

Для минимизирования SSR_q необходимо заранее для каждого рассматриваемого момента t_i определить вектор $\underline{r}(t_i)$ с помощью NNLS-алгоритма таким образом, чтобы

$$\min_{\underline{r}(t_i)} (\|L \cdot \underline{r}(t_i) - \tilde{q}(t_i)\|) \quad (\text{формула 13})$$

с дополнительным краевым условием

$$\underline{r}(t_i) \geq 0 \quad (\text{формула 14})$$

Преимуществом этого метода является то, что также можно проводить расчеты по формулам 12-14 для большого количества подмножеств со многими ЕМ.

Значительным недостатком является то, что для этих расчетов требуются рассчитанные специфические нормы $\tilde{q}(t)$. Так как они получены из интерполированных данных измерений, то между ними и их действительными значениями лежит значительная разница. Неточность измерений в определенных обстоятельствах может оказывать сильное влияние на рассчитанные специфические нормы $\tilde{q}(t)$. Следовательно, критерий качества SSR_q можно определить с большой неточностью. Кроме информации о качестве изображения с помощью этого метода также получают оцененное направление скорости реакции $\underline{r}(t)$ подмножества L в качестве результата минимизации по формулам 13 и 14.

Leighty, R. и др. описывает следующий метод, в котором данные измерений (измерений концентрации) сразу приближают с помощью линейной оценки волюметрических скоростей реакций через промежутки времени. При решении линейной проблемы оптимизации с помощью решающей программы наименьших квадратов можно быстро оценить направление реакции при условии, что оно проходит линейно между опорных точек [Leighty, R.W., & Antoniewicz, M.R. (2011), Dynamic metabolic flux analysis (DMFA): a framework for determining fluxes at metabolic non-steady state. Metabolic engineering, 13(6), 745-755]. Этот метод относится только к обратимым макрореакциям (как их называют в источнике "Free Fluxes"), кроме того, можно не учитывать эффекты разбавления (т.е. изменения концентраций, которые не вызваны клетками). Если размер макрореакций и данные измерений не совпадают, то эти данные измерений нельзя использовать для оценивания скоростей реакций. Это известно, например, в случае, когда рост клеток в виде образования внешней биомассы является частью макрореакций и известны данные измерений только для сухой массы клеток. Поэтому в этом виде этот метод не подходит для применения в необратимых макрореакциях и процессах с подпиткой.

Если применяют концепт Leighty и др. и его по отзывам добавляют к этой заявке с предварительно обработанными согласно изобретению (со сдвигом) данными, то этот метод также можно применять для процессов с подпиткой. Кроме того, с помощью дополнения нижней границы для скоростей макрореак-

ции в виде краевого условия линейной проблемы оптимизации можно также применять метод для необратимых реакций, как элементарные моды. Если размер макрореакций и данные измерений не совпадают, то с помощью подходящей корреляции можно согласовать размер данных измерения с данными макрореакций. Эту комбинацию линейного оценивания согласно Leighty и др. с улучшениями по этому пункту называют далее "линейное оценивание скоростей реакций выбранных макрореакций".

Необходимо проверить, можно ли с выбранными макрореакциями подмножества L оригинального ЕМ-комплекта К предоставить определенные данные в достаточной мере. Рассчитанная здесь финальная сумма возведенных в квадрат SSR_C по формуле 15 между определенными с помощью метода концентра-

циями $\hat{C}_s(t)$ со сдвигом и концентрациями $C_{js}(t)$ со сдвигом показывает, насколько хорошо данные измерений можно отобразить с подмножеством.

$$SSR_C = \sum_{i=1}^{N_t} \sum_{j=1}^{N_c} (C_{js}(t_i) - \hat{C}_{js}(t_i))^2 \quad (\text{формула 15})$$

Чем меньше значение SSR_C , тем лучше подмножество L. Этот метод, по сравнению с методом Soons и др. является более предпочтительным, особенно для моделирования процессов с подпиткой, так как можно осуществить быструю проверку качества подмножества без возможно ошибочного предыдущего определения специфических норм. При условии, что оцениваемые скорости реакции проходят линейно между точками опоры, погрешности измерений оказывают лишь незначительное воздействие на оценку скоростей реакций. Недостатком этого метода является то, что размер исследуемого подмножества L ограничен решением линейной проблемы оптимизации. Максимальное количество реакций в подмножестве равно количеству проведенных измерений, разделенных на количество точек опоры.

Кроме информации о качестве изображения с помощью этого метода также получают оцененное направление скоростей реакции подмножества $\underline{r}(t)$.

В предпочтительной форме выполнения "линейного оценивания скоростей реакций выбранных макрореакций" согласно изобретению для решения линейной проблемы оптимизации вместо LLSS используют NNLS по Lawson и др. [Lawson, C.L. and R.J. Hanson, Solving Least Squares Problems, Prentice-Hall, 1974, Chapter 23, стр. 161]. С помощью этого также можно перепроверить качество больших подмножеств с помощью способа. При этом максимальное количество макрореакций значительно больше, чем количество проведенных измерений, разделенных на количество точек опоры. Эту комбинацию "линейного оценивания скоростей реакций выбранных макрореакций" с применением NNLS называют далее "линейное оценивание скоростей реакций выбранных макрореакций с NNLS".

Метод согласно изобретению "линейного оценивания скоростей реакций выбранных макрореакций с NNLS" можно дополнительно использовать в качестве другого зависимого от данных метода предварительного сокращения ЕМ на этапе е) i). Для этого здесь можно использовать большой набор К макрореакций. Результатом этого метода, во-первых, является значение для SSR_C и, во-вторых, направление скоростей реакций $\underline{r}(t)$. ЕМ с маленькими значениями принадлежащих норм $\underline{r}(t)$ удаляют из матрицы К. Этот способ действия повторяют, пока не достигнут заранее определенного количества ЕМ или не превысят значение SSR_C установленного предельного значения.

Алгоритмы выбора подмножества известны, например, из Provost и др. и Soons и др. [Provost A. 2006. Metabolic design of dynamic bioreaction models. Faculté des Sciences Appliquées, Université catholique de Louvain, Louvain-la-Neuve, Louvain-la-Neuve; Soons, Z.I.T.A., Ferreira, E.C., Rocha, I. (2010). Selection of Elementary Modes for Bioprocess Control. 11th International Symposium on Computer Applications in Biotechnology, Левен, Бельгия, июль 7-9, 2010, 156-161].

Soons и др. описывают образование ЕМ-подмножества в двухступенчатом методе оптимизации. Для различных, случайно выбранных ЕМ значения для SSR_q , как описано выше, соответственно минимизируют. Выбирают набор с наименьшим минимизированным SSR_q значением. Однако при большом числе ЕМ случайный выбор является неэффективным, так как количество возможных комбинаций сильно растет. Например, при выборе 10 реакций из комплекта 20.000 ЕМ есть более чем $2,8 \cdot 10^{36}$ комбинаций. Вероятность, что будет найдена оптимальная комбинация, очень мала. При использовании критерия качества SSR_q этот метод становится очень чувствительным по отношению к погрешностям и отклонениям в измерениях.

Provost описывает альтернативный алгоритм, в котором для различных специфических значений $\underline{\tilde{q}}(t_i)$ $i = 1, \dots, n$ устанавливают все возможные положительные линейные комбинации элементарных мод, в которых $SSR_q(t=t_i)=0$. Затем из этих многочисленных возможных комбинаций случайно выбирают одну. Соответственно этот способ применяет только вектор $\underline{\tilde{q}}(t_i)$ и не все направление в течение всего времени. Поэтому выбор ЕМ для всего процесса невозможен. Однако с помощью случайного выбора можно представить вектор $\underline{\tilde{q}}(t_i)$, насколько можно этим представить остаток процесса. Другим недостатком этого способа является то, что вектор $\underline{\tilde{q}}(t_i)$, который не находится в пространстве решений всех

ЕМ, нельзя использовать. Нельзя определить приближенное решение. Это является большим недостатком, прежде всего при недостоверности имеющихся измерений и специфических норм. При использовании критерия качества SSR_q этот метод также становится очень чувствительным по отношению к погрешностям и отклонениям в измерениях.

Поэтому в другой предпочтительной форме способа для выбора относительных макрореакций, т.е. выбора ЕМ-подмножества L применяют эволюционированный, в частности, генетический алгоритм. Такой алгоритм известен, например, из Baker и др. [Syed Murtuza Baker, Kai Schallau, Björn H. Junker. 2010. Comparison of different algorithms for simultaneous estimation of multiple parameters in kinetic metabolic models. J. Integrative Bioinformatics:-1-1.]. Особенно предпочтительно можно применять генетический алгоритм, в котором рассчитывают целевую функцию для различных комбинаций ЕМ соответствующего значения SSR_c с помощью метода "линейного оценивания скоростей реакций выбранных макрореакций". Также альтернативно можно использовать случайный выбор. После окончания этапа ii) матрица L содержит необходимые макрореакции (этап iii).

На проводимом, при необходимости, этапе iii) графически проверяют достоверность ЕМ-подмножества L. Здесь можно использовать Flux Map из этапа d) в виде проекции ЕМ-подмножества L. Фиг. 4 показывает Flux Map с проекцией подмножества шести ЕМ. Если ЕМ-подмножество L является действующим, то данные измерений находятся внутри ЕМ-подмножества L. Это изображение дает возможность быстрого графического контроля над действенностью выбора.

На следующем этапе f) с помощью специфических норм $\frac{\tilde{q}(t)}{\tilde{q}(t)}$ из этапа d) и/или из данных измерений из этапа c) рассчитывают данные макрореакций подмножества L. Расчет $\tilde{q}(t)$ можно проводить на основании специфических норм $\frac{\tilde{q}(t)}{\tilde{q}(t)}$, как описано в этапе e) по Soons и др. [Soons, Z. I. T. A., Ferreira, E. C., Rocha, I. (2010). Selection of Elementary Modes for Bioprocess Control. 11th International Symposium on Computer Applications in Biotechnology, Leuven, Бельгия, июль 7-9, 2010, 156-161], предпочтительно расчет $\tilde{q}(t)$ проводят на основании данных измерений из этапа c) с помощью "линейного оценивания скоростей реакций выбранных макрореакций" согласно изобретению.

На этапе g) способа проектируют кинетики макрореакций. Установленные кинетики должны дать количественную оценку динамическому воздействию состояния процесса на соответствующие скорости реакции $\hat{r}_k$:

$$\hat{r}_k = f(C, pH, T, \dots) \quad (\text{формула 16})$$

Из кинетик определяют определяющие параметры модели. Родовые кинетики на этапе g) i) проектируют из стехиометрии макрореакций. Для субстрата макрореакций i принимают во внимание лимитирование Monod типа.

Максимальную скорость реакции умножают на различные лимитирования $\tilde{r}_i$:

$$\hat{r}_k(t) = r_{k,max} \cdot \prod_{i=1}^{N_l} \tilde{r}_i(t) = r_{k,max} \cdot \prod_{i=1}^{N_l} \left(\frac{C_i(t)}{K_{m,k,i} + C_i(t)} \right)^{n_i} \quad (\text{формула 17})$$

Причем постоянные Моно $K_{m,k,i}$ и коэффициент Хилла n_i представляют параметр уравнения, первое значение которого вводят вручную. Обычно вводят постоянные Моно $K_{m,k,i}$ на одну десятую часть соответствующих максимально определенных концентраций и коэффициенты Хилла n_i на значение 1. Определение видовых кинетик из стехиометрии реакции описали Provost, Gao и др. [Provost A. 2006. Metabolic design of dynamic bioreaction models. Faculté des Sciences Appliquées, Université catholique de Louvain, Louvain-la-Neuve, S. 126; [Gao, J., Gorenflo, V. M., Scharer, J. M., & Budman, H. M. (2007). Dynamic metabolic modeling for a MAB bioprocess. Biotechnology progress, 23(1), 168-181]. По отзывам эти методы интегрируют в заявку. В этих методах применяют лимитирование субстрата типа Моно для соответствующих субстратов реакции. Хотя Provost или Gao это не описывают, ингибирования также токсичными продуктами с помощью этого метода вытекают из стехиометрии реакции.

На этапе g) ii) устанавливают параметры влияния на установленные на этапе f) скорости реакций $\tilde{q}(t)$. При этом рассматривают все величины, которые описывают состояние процесса (т.е. такие условия биореакции как, например, значение pH, температуру реактора, парциальное давление, которые не исходят из стехиометрии макрореакции). Параметры воздействия можно установить вручную, например с помощью статистического метода, как метод частных наименьших квадратов (PLS). Кроме того, устанавливают корреляцию между состоянием процесса (которое присутствует в матрице) и скоростями реакций $\tilde{q}(t)$ из этапа f).

Затем на этапе g) iii) дают количественную оценку установленных на этапе g) ii) воздействий и дополняют кинетику из i) соответствующими элементами. Обнаруженные на этапе g) ii) параметры воздействия состояния процесса на скорость реакции можно подтвердить элементом $\tilde{r}_i$. Элемент $\tilde{r}_i$ является любой функцией, которая в зависимости от состояния процесса обнаруживает значение 0-1. Установлен-

ная на этапе g) i) родовая кинетика реакции умножается затем на этот элемент.

Например, обнаруженная негативная корреляция между концентрацией компонента i и реакцией k указывает на влияние скорости реакции k с помощью концентрации i (C_i). Это можно доказать, например, с помощью кинетики ингибирования по Холдейну:

$$\hat{r}_i(t) = \frac{K_{i,k,i}}{K_{i,k,i} + C_i(t)} \quad (\text{формула 18})$$

Причем $K_{i,k,i}$ обозначает константу ингибирования и является другим параметром модели, первое значение которой вводят вручную, и обычно составляет десятую часть соответствующей максимальной определенной концентрации.

На проводимом при необходимости этапе h) значения параметров модели $\underline{p}$ кинетик согласовывают с установленными на этапе f) скоростями реакций макрореакций $\underline{r}(t)$:

$$\min_{\underline{p}} \sum_{k=1}^{N_t} (\hat{r}_k(\underline{p}) - r_k)^2 \quad (\text{формула 19})$$

Далее это называют оценкой параметров модели. От числового решения одного или более дифференциальных уравнений по формулам 2 б-4 на этом этапе можно отказаться; значения параметров модели можно согласовать в независимых группах обычно с 3-10 параметрами отдельно для каждой макрореакции k .

Согласование проводят с помощью обычного метода как метод Гаусса-Ньютона [Bates DM, Watts DG. 1988. Nonlinear regression analysis and its applications. Нью-Йорк: Wiley. xiv, 365].

Это отдельное для каждой макрореакции оценка параметров модели является особенно предпочтительной для этапов i) и j), так как ее, с одной стороны, можно быстро провести и, с другой стороны, она дает улучшенные начальные значения для согласования значений параметров модели с данными измерений из этапа c) на этапе j).

Качество согласования рассчитывают, например, с суммой возведенных в квадрат остатков SSR_r по формуле 20:

$$SSR_r = \sum_{k=1}^{N_t} (\hat{r}_k(\underline{p}) - r_k)^2 \quad (\text{формула 20})$$

Чем меньше значение для SSR_r , тем лучше согласование. Альтернативно проводят перепроверку качества согласования с помощью графического сравнения $\hat{r}_k$ и r_k .

На проводимом при необходимости этапе i) перепроверяют выбранные на этапе g) кинетики макрореакций на качество их согласования. Основой является рассчитанное на этапе h) значение SSR_r , которое дает количественную оценку качеству согласования оценки параметров модели. При неудовлетворительном качестве согласования можно повторить этапы g) и h) до получения заданного согласования.

На следующем этапе j) можно провести согласование значений параметров кинетик из этапа g) с данными измерений из этапа c) согласно одному из методов, обычных для согласования. Предпочтительно для этого согласования применяют начальные параметры из этапа h). Согласование значений параметров моделей проводят с применением вышеуказанного дифференциального уравнения (формулы 2-4), например с помощью метода Ньютона-Гаусса [Bates D.M., Watts D.G. 1988. Nonlinear regression analysis and its applications. Нью-Йорк: Wiley. xiv, 365] или с помощью multiple Shooting алгоритма [Peifer M., Timmer J. 2007. Parameter estimation in ordinary differential equations for biochemical processes using the method of multiple shooting. Systems Biology, IET 1(2):78-88].

Предпочтительно в модель можно интегрировать признаки продукта. Особенно предпочтительно это можно проводить, если признаки зависят от концентрации побочных и промежуточных продуктов. Концентрации побочных продуктов, внешние компоненты введенной в а) метаболической сети уже интегрированы в модель и их можно рассчитать. При необходимости также можно объединять другие побочные и промежуточные продукты в одну или несколько отдельных метаболических сетей. Это является предпочтительным, когда ожидаемые нормы выхода и поступления принадлежат к разному порядку величин, или необходимо рассматривать определенные процессы обмена веществ в разных детальных степенях. Альтернативно к интегрированной модели можно получить отдельную модель для расчета признаков продукта с этапами а)-j), так что протекание процесса внешних компонентов отдельной метаболической сети также описывают одним предложением о макрореакциях с собственной кинетикой. Побочные и промежуточные продукты, которые не находятся вне организма, но внутриклеточная аккумуляция которых может воздействовать на один или несколько признаков продукта, экстернализируют на этапе а) и б) при расчете ЕМ и композиции макрореакций, т.е. классифицируют как внешние компоненты. Связь признаков продукта, которые зависят от внутриклеточных или внеклеточных концентраций, может осуществляться через дополнительную интеграцию количественных или качественных связей между концентрациями и признаками продукта.

Другими предметами изобретения является компьютерная программа или ПО для проведения способа согласно изобретению.

Полученную с помощью способа согласно изобретению модель можно применять для технологического процесса или планирования производственного процесса, а также для анализа способа в реакторе.

Особая форма выполнения способа согласно изобретению описана только в качестве примера, ею не стоит ограничиваться. С помощью этого способа также в виде примера получили модель ферментации с клетками гибридомы и проверили его действенность, как было описано.

Пример. Моделирование культуры клеток гибридомы.

1. Этап а).

Базовые сведения в виде метаболической сети взяли из публикации Niu и др. (Metabolic pathway analysis and reduction for mammalian cell cultures-Towards macroscopic modeling. Chemical Engineering Science (2013) 102, стр. 461-173. DOI: 10.1016/j.ces.2013.07.034). Описанная здесь метаболическая сеть клетки животного содержит 35 реакций, которые связывают 37 внутренних и внешних метаболитов (см. фиг. 5, см. табл. 1).

Таблица 1. Реакции метаболической сети по Niu и др. (Metabolic pathway analysis and reduction for mammalian cell cultures-Towards macroscopic modeling. Chemical Engineering Science (2013) 102, стр. 461-473. DOI: 10.1016/j.ces.2013.07.034.)

1 глюкоза $\rightarrow$ 1 G6P

1 G6P + 2 NAD $\rightarrow$ 2 пируват

1 пируват $\rightarrow$ 1 лактат + 1 NAD

1 пируват $\rightarrow$ 1 пируват_m

1 NADm + 1 пируват_m $\rightarrow$ 1 ацетил coA_m

1 ацетил coA_m + 1 NADm + 1 оксалоацетат_m $\rightarrow$ 1 α -кетоглутарат_m

1 -кетоглутарат_m + 1 NADm $\rightarrow$ 1 сукцинил CoA_m

1 FADm + 1 сукцинил CoA_m → 1 фумарат
 1 фумарат → 1 малат_m
 1 малат_m + 1 NADm → 1 оксалоацетат_m
 1 глютамин → 1 глютамат + 1 NH₃
 1 глютамат + 1 NADm → 1 α-кетоглутарат_m + 1 NH₃
 1 малат_m → 1 малат
 1 малат + 1 NAD → 1 пируват
 1 глютамат + 1 пируват → 1 α-кетоглутарат_m + 1 аланин
 1 глютамат + 1 оксалоацетат_m → 1 α-кетоглутарат_m + 1 аспартат
 1 аргинин + 2 NADm → 1 глютамат + 3 NH₃
 1 аспарагин → 1 аспартат + 1 NH₃
 2 глицин + 1 NADm → 1 NH₃
 1 гистидин + 1 NADm → 1 глютамат + 2 NH₃
 1 изолейцин + 2 NADm → 1 ацетил coA_m + 1 NH₃ + 1 сукцинил CoA_m
 1 лейцин + 3 NADm → 3 ацетил coA_m
 1 лизин + 6 NADm → 2 ацетил coA_m
 1 метионин + 4 NADm → 1 NH₃ + 1 сукцинил CoA_m
 1 NADm + 1 феналанин → 1 тирозин
 1 серин → 1 NH₃ + 1 пируват
 1 NADm + 1 треонин → 1 NH₃ + 1 сукцинил CoA_m
 19 NADm + 1 TRP → 3 ацетил coA_m
 5 NADm + 1 тирозин → 2 ацетил coA_m + 1 фумарат
 5 NADm + 1 валин → 1 NH₃ + 1 сукцинил CoA_m

1 NADm → 1 NAD

0.5 кислород(O₂) → 1 NADm

1 NADm → 1 FADm

0.0156 аланин + 0.0082 аргинин + 0.0287 аспартат + 0.0167 G6P + 0.0245 глутамин +
0.0039 глутамат + 0.0196 глицин + 0.0038 гистидин + 0.0099 изолейцин + 0.0156
лейцин + 0.0119 лизин + 0.0039 метионин + 0.0065 феналанин + 0.016 серин + 0.0094
треонин + 0.0047 тирозин + 0.0113 валин → 1 X (**биомасса**) + 0.0981 NAD

0.01101 аланин + 0.005033 аргинин + 0.007235 Asparagine + 0.0081787 аспартат +
0.010381 глутамин + 0.010695 глутамат + 0.01447 глицин + 0.0034602 гистидин +
0.005033 изолейцин + 0.014155 лейцин + 0.01447 лизин + 0.0028311 метионин +
0.007235 феналанин + 0.026738 серин + 0.016043 треонин + 0.0084932 тирозин +
0.018874 валин → 1 IgG (**антитело**)

В заявке обратимость реакций указана неопределенно. Вместо этого предоставлены данные для "метаболического анализа потока" из этой же самой заявки, и они использовались для идентификации необратимых реакций.

С помощью стехиометрической матрицы N, которая содержит стехиометрию, т.е. стехиометрические коэффициенты внутренних метаболитов, и информации о обратимости реакций, также с использование METATOOL 5.1 (Pfeiffer et al. METATOOL: for studying metabolic networks, Bioinformatics 199915 (3), S. 251-257) рассчитали все элементарные моды (ЕМ) сети. Число ЕМ здесь составляет более 300.000.

2. Этап b).

Матрицу с рассчитанными ЕМ E получили на этапе а). Аналогично матрице N получили матрицу N_p стехиометрии, т.е. стехиометрические коэффициенты, внешний метаболит. Возможные макрореакции стехиометрической сети получили в матрице K с формулой 21:

$$K = N_p \cdot E \quad (\text{формула 21})$$

3. Этап c).

Данные измерений процесса взяли из Vaughman и др., который указал различные параметры ферментации клеток гибридомы в течение процесса пакетной обработки (см. фиг. 6) [On the dynamic modeling of mammalian cell metabolism and mAb production. In: Computers & Chemical Engineering (2010) 34 (2), стр. 210-222]. Данные измерений ввели в способ.

4. Этап d).

С помощью сплайн-интерполированных данных измерений из этапа c) ($\underline{C}^{int}$) рассчитали скорость роста и показатель гибели микроорганизмов, а также специфические скорости поступления или выхода (см. фиг. 7). Лизис включили с заранее определенным и введенным в способ через определенный промежуток времени постоянным лизис-фактором K_l=0.1. Сдвиг данных измерения не требовался, так как здесь речь идет о данных процесса пакетной обработки без последующих добавлений. Соответственно этому данные показывают усиление процесса, так как изменения концентраций вызваны клетками, а не добавлением.

Для расчета норм $\frac{\tilde{q}}{l}$ использовали дополнительную информацию. Так с помощью, названной в записях из Vaughman и др. общей биомассы (BM в $\left[\frac{C-mol}{l}\right]$) и общего количества клеток смогли рассчитать среднее C-моль содержание $f_{C-molX_p} = 18.41 \left[\frac{C-mol}{10^9 cells}\right]$. Относящийся к C-моль скорость роста можно рассчитать только с помощью формулы 22:

$$\tilde{\mu} \left[\frac{C-mol}{h \cdot 10^9 cells} \right] = \mu \left[\frac{1}{h} \right] \cdot f_{C-molX_p} \left[\frac{C-mol}{10^9 cells} \right] \quad (\text{формула 22})$$

Аналогично можно оценить C-моль относительную норму образования антитела. Кроме того, оценили молярный состав антитела к CH_{1.58}O_{0.31}N_{0.27}S_{0.004} с формальной молярной массой M_{mAb,C-mol} = 22.45 $\frac{g}{mol}$. Считают, что молярный состав соответствует среднему молярному составу протеинов, который указан в Villadsen и др. [Bioreaction engineering principles (2011), глава 3, Elemental and Redox Balances, S. 73, изд. Springer Verlag, ISBN: 978-1-4419-9687-9]. Молярную массу всего антитела оценили, как M_{mAb} = 150.000 $\frac{g}{mol}$. Норма образования антител получилась из формулы

$$\tilde{q}_{mAb} \left[\frac{C - mol}{h \cdot 10^9 cells} \right] = q_{mAb} \left[\frac{10^{-4} mol}{h \cdot 10^9 cells} \right] \cdot \frac{M_{mAb, C-mol}}{M_{mAb}} \cdot 1 \quad (\text{формула 23})$$

Изменение норм во времени $\tilde{q}(t)$ можно было использовать для выбора макрореакций.

5. Этап е).

На этапе е) получили ЕМ-подмножество макрореакций, с которым наилучшим образом передали набор данных. Для этого использовали матрицу К из этапа б). Т.к. количество более 300.000 макрореакций могло бы привести к большому количеству возможных комбинаций, то сначала провели зависимое от данных предварительное сокращение.

Для этого использовали установленные на этапе d) нормы $\tilde{q}(t)$ для расчета коэффициентов выхода продукции Y^m для всех комбинаций двух внешних метаболитов. Нижнюю границу коэффициента выхода продукции Y_{ij} выбрали таким образом, что 99% установленных коэффициентов выхода продукции $Y_{ij}^m(t)$ лежали над этим значением. Верхнюю границу коэффициента выхода продукции выбрали таким образом, что 99% установленных коэффициентов выхода продукции $Y_{ij}^m(t)$ лежали ниже этого значения. В качестве примеров указаны некоторые установленные границы, а также количество ЕМ, коэффициенты выхода продукции Y_{ij}^{EM} которых находятся внутри этих границ.

Табл. 2: итог смогли сократить ЕМ примерно до 3000.

Таблица 2. Внешние метаболиты, максимальные и минимальные коэффициенты выхода продукции Y_{ij} которых, а также количество ЕМ, коэффициенты выхода продукции которых находятся внутри указанных границ

внешние компоненты	нижняя граница	верхняя граница	количество ЕМ внутри границ
Ala : Asn	-14.6811	-0.1752	99.9134 %
Asn : Glc	0.0293	16.1603	64.0488 %
Asp : Ala	-1.1130	-2.1011	74.0839 %

Затем вслед за зависимым от данных сокращением провели еще независимую от данных редукцию. При этом определили максимальное значение для косинус-подобий двух ЕМ 0,995. Начиная с первой реакции удалили все макрореакции из матрицы К, которые превышали это значение. Осталось примерно 500 макрореакций из матрицы К (названной также приведенная матрица К), которые заняли затем более 95% объема пространства решений примерно с 3000 ЕМ.

Кроме того, перед процессом выбора провели выравнивание указанных в метаболической сети по Niu и др. компонентов (которые соответствуют внешним метаболитам метаболической сети из этапа а)) и обнаруженных концентраций компонентов из этапа с). До пролина принимают во внимание все определенные Vaughman и др. концентрации, также в метаболической сети по Niu и др. Для использования измерения концентрации пролина можно было бы или применить другую упрощенную сеть, которая содержит пролин в виде внешних метаболитов, или провести расширение существующей метаболической сети.

Компоненты, которые встречались в оцениваемых макрореакциях, но данными которых нельзя было воспользоваться, также далее игнорировали. Соответствующие линии матрицы К соответственно этому из нее вычеркивали.

Вычеркивание соответствующих линий не означает, что их нельзя использовать. Они состоят в метаболической сети, только отсутствуют измерения, которые смогли бы их уравновесить. В этом примере не уделили внимания поступлению и выходу аргинина, глутамата, глицина, гистидина, лейцина, лизина, метионина, аммония, кислорода, фенилаланина, серина, треонина, триптофана, тирозина и валина.

На следующих этапах способа использовали только приведенную матрицу К, которая представляет базовые сведения, и данные $\tilde{q}(t)$ из этапа d), а также данные измерений из этапа с), которые образуют технологические знания для того, чтобы получить наименьшее подмножество L макрореакций из К.

В качестве критерия качества использовали "линейную оценку скоростей реакций, выбранных макрореакций" согласно изобретению.

Аналогично нормам $\tilde{q}(t)$ в относительных величинах выразили данные измерений количества клеток и антитела до С-моль. Это требуется для того, чтобы величины макрореакций совпадали с величинами данных измерений.

Выбор подмножества провели с генетическим алгоритмом. При этом в вычислении целевой функции этого генетического алгоритма решили рассматриваемую в "линейной оценке скоростей реакций, избранных макрореакций" линейную проблему оптимизации. Рассчитанная здесь финальная сумма самой маленькой квадратичной погрешности линейной проблемы оптимизации была равна значению целе-

вой функции для соответствующего выбора макрореакций.

Для выбора размера подмножества L из K несколько раз провели оптимизацию с различным количеством макрореакций в L. Количество является компромиссом между комплексностью модели и точностью передачи. Для того чтобы установить, сколько реакций хватает для передачи, можно или повторить выбор подмножества L для разного количества макрореакций, или сразу в целевую функцию генетического алгоритма добавить дополнительный элемент для количества реакций. В этом случае проводили несколько оптимизаций с заранее определенным количеством макрореакций (10, 7, 5, 4 и 3). Обнаруженные с помощью генетического алгоритма самые маленькие суммы квадратичной погрешности представлены на фиг. 9 напротив количества макрореакций. Оказалось, что в этом случае меньше семи макрореакций является недостаточным, чтобы представить ход технологического процесса. Выбранные реакции указаны в табл. 3.

Таблица 3. Выбранное подмножество макрореакций (L)
(неподчеркнутые компоненты в модели не принимали во внимание,
так как для них не было измерений)

0.474 аланин + 0.474 метионин

→ 0.158 аспарагин + 0.316 аспартам + 0.632 глицин
+ 0.158 триптофан

0.015 аланин + 0.00789 аргинин + 0.304 аспарагин + 0.0161 глюкоза
+ 0.0236 глутамин + 0.00375 глутамат + 0.00366 гистидин
+ 0.00953 изолейцин + 0.015 лейцин + 0.112 метионин
+ 0.00626 феналанин + 0.0154 серин + 0.0109 валин
→ 0.963 X (биомасса) + 0.00276 аспартам + 0.24 глицин
+ 0.0208 триптофан

0.295 аспарагин + 0.147 глутамат
+ 0.295 аспартам + 0.885 глицин + 0.147 лактат

0.00753 аргинин + 0.113 аспарагин + 0.0603 глюкоза + 0.0225 глутамин
+ 0.0824 гистидин + 0.00909 изолейцин + 0.00597 феналанин
+ 0.0216 триптофан + 0.00431 тирозин + 0.0104 валин
→ 0.918 X (биомасса) + 0.061 аланин + 0.0865 аспартам
+ 0.343 глицин + 0.0631 метионин

0.0654 аргинин + 0.412 аспартам + 0.00991 глюкоза + 0.0145 глутамин
+ 0.554 глицин + 0.00226 гистидин + 0.00588 изолейцин
+ 0.00926 лейцин + 0.00706 лизин + 0.0649 феналанин
+ 0.0095 серин + 0.0671 валин
→ 0.594 X (биомасса) + 0.049 аланин + 0.395 аспарагин
+ 0.0503 треонин + 0.0388 триптофан

$$\begin{aligned}
& 0.0077 \text{ аргинин} + \underline{0.179 \text{ аспарат}} + \underline{0.0157 \text{ глюкоза}} + \underline{0.104 \text{ глутамин}} \\
& + 0.216 \text{ глицин} + 0.00357 \text{ гистидин} + 0.00929 \text{ изолейцин} \\
& + 0.0146 \text{ лейцин} + 0.0112 \text{ лизин} + 0.038 \text{ тирозин} + 0.0106 \text{ валин} \\
& \rightarrow \underline{0.939 X (\text{биомасса})} + \underline{0.0624 \text{ аланин}} + \underline{0.152 \text{ аспарагин}} \\
& + 0.0183 \text{ триптофан}
\end{aligned}$$

$$\begin{aligned}
& 0.0342 \text{ аргинин} + \underline{0.211 \text{ аспарат}} + \underline{0.00762 \text{ глюкоза}} + \underline{0.0195 \text{ глутамин}} \\
& + 0.244 \text{ глицин} + 0.00452 \text{ гистидин} + 0.0546 \text{ изолейцин} \\
& + 0.0185 \text{ лейцин} + 0.0171 \text{ лизин} + 0.00406 \text{ метионин} \\
& + 0.0178 \text{ тирозин} + 0.0203 \text{ валин} \\
& \rightarrow \underline{0.457 X (\text{биомасса})} + \underline{0.804 \text{ IgG (антитело)}} + \underline{0.185 \text{ аспарагин}} \\
& + 0.0153 \text{ триптофан}
\end{aligned}$$

В указанных макрореакциях названы все внешние метаболиты метаболической сети из этапа а). Однако частью модели являются только подчеркнутые внешние метаболиты, так как только для них присутствуют данные измерений из этапа с).

6. Этап f).

Для выбранного комплекта макрореакций установили скорости реакций через период времени. В этом примере с помощью метода "линейная оценка скоростей реакций избранных макрореакций" согласно изобретению приблизили показанные на фиг. 10 данные измерений с помощью оценивания скоростей реакций $\hat{r}(t)$.

Результатом метода является кусочно-линейное направление отдельных (волюметрических) скоростей реакций. С помощью деления на интерполированное направление количества жизнеспособных клеток $X_v(t)$ получили особые для клеток скорости реакций $\hat{r}(t)$, показанных в табл. 3 макрореакций. Полученные таким образом скорости реакций $\hat{r}(t)$ представлены на фиг. 10.

7. Этап g).

Для всех показанных в табл. 3 макрореакциях принимали во внимание родовые кинетики согласно формулы 24:

$$\hat{r}_k(t) = r_{k,max} \cdot \prod_{i=1}^{N_l} \hat{r}_i(\underline{c}(t), \underline{p}, \dots) \quad (\text{формула 24})$$

В этом случае их выполняют с помощью кинетики Жака Моно. Т.е. во время каждой реакции k для каждого субстрата i вводили лимитирование согласно формуле 25:

$$\hat{r}_i(t) = \left(\frac{C_i(t)}{K_{m,k,i} + C_i(t)} \right)^{n_i} \quad (\text{формула 25})$$

При этом $r_{k,max}$ отображает максимальную скорость реакции, N_l количество учитываемых ограничений, C_i концентрацию компонента i , $K_{m,k,i}$ относящиеся к этому постоянные Моно и n_i параметр Хилла для порядка реакции. Их значения согласовывают на этапах h) и j).

Другие элементы получаются из анализа скоростей реакции $\hat{r}(t)$ из этапа f). В этом примере кроме лимитирований субстрата также учитывают ингибирования согласно формуле 26:

$$\hat{r}_i = \left(\frac{K_{I,k,i}}{K_{I,k,i} + C_i} \right)^{n_i} \quad (\text{формула 26})$$

Для этого лимитирования также необходимо согласовать значения параметров $K_{I,k,i}$ и n_i . Применяемые кинетические элементы реакций представлены в табл. 4.

Таблица 4. Кинетические элементы выбранных макрореакций из L

$$\hat{r}_1(t) = r_{1,max} \cdot \left(\frac{[Ala](t)}{K_{m,Ala,1} + [Ala](t)} \right) \cdot \left(\frac{[Glc](t)}{K_{m,Glc,1} + [Glc](t)} \right)^2 \cdot \left(\frac{K_{I,Asn,1}}{K_{I,Asn,1} + [Asn](t)} \right)^2$$

$$\hat{r}_2(t) = r_{2,max} \cdot \left(\frac{[Glc](t)}{K_{m,Glc,2} + [Glc](t)} \right) \cdot \left(\frac{[Gln](t)}{K_{m,Gln,2} + [Gln](t)} \right) \cdot \left(\frac{[Asn](t)}{K_{m,Asn,2} + [Asn](t)} \right) \cdot \left(\frac{[Ala](t)}{K_{m,Ala,2} + [Ala](t)} \right)^2$$

$$\hat{r}_3(t) = r_{3,max} \cdot \left(\frac{[Glc](t)}{K_{m,Glc,3} + [Glc](t)} \right) \cdot \left(\frac{[Asn](t)}{K_{m,Gln,3} + [Asn](t)} \right) \cdot \left(\frac{K_{I,Lac,3}}{K_{I,Lac,3} + [Lac](t)} \right)^2$$

$$\hat{r}_4(t) = r_{4,max} \cdot \left(\frac{[Glc](t)}{K_{m,Glc,4} + [Glc](t)} \right) \cdot \left(\frac{[Gln](t)}{K_{m,Gln,4} + [Gln](t)} \right) \cdot \left(\frac{[Asn](t)}{K_{m,Asn,4} + [Asn](t)} \right) \cdot \left(\frac{[X_t](t)}{K_{m,Xt,4} + [X_t](t)} \right)$$

$$\hat{r}_5(t) = r_{5,max} \cdot \left(\frac{[Glc](t)}{K_{m,Glc,5} + [Glc](t)} \right) \cdot \left(\frac{[Gln](t)}{K_{m,Gln,5} + [Gln](t)} \right) \cdot \left(\frac{[Asp](t)}{K_{m,Asp,5} + [Asp](t)} \right) \cdot \left(\frac{K_{I,Asp,5}}{K_{I,Asp,5} + [Asp](t)} \right)^3 \cdot \left(\frac{K_{I,Asn,5}}{K_{I,Asn,5} + [Asn](t)} \right)$$

$$\hat{r}_6(t) = r_{6,max} \cdot \left(\frac{[Glc](t)}{K_{m,Glc,6} + [Glc](t)} \right) \cdot \left(\frac{[Gln](t)}{K_{m,Gln,6} + [Gln](t)} \right) \cdot \left(\frac{[Asp](t)}{K_{m,Asp,6} + [Asp](t)} \right)^2$$

$$\hat{r}_7(t) = r_{7,max} \cdot \left(\frac{[Glc](t)}{K_{m,Glc,7} + [Glc](t)} \right) \cdot \left(\frac{[Gln](t)}{K_{m,Gln,7} + [Gln](t)} \right) \cdot \left(\frac{[Asp](t)}{K_{m,Asp,7} + [Asp](t)} \right)$$

8. Этап h).

Для каждой скорости реакции с помощью указанной в табл. 4 кинетики и интерполированных значений учитываемых к кинетике концентраций $\underline{C}^{int}(t)$ смогли алгебраически рассчитать направление скоростей реакции

$$\hat{r}_i(\underline{p}, \underline{C}^{int}(t))$$

Согласование параметров этой кинетики проводят отдельно для каждой реакции i с установленной на этапе f) скоростью реакции $r_i(t)$. Целевая функция для оптимизации встречающихся в реакции i параметров в этом примере следующая:

$$\min_{\underline{p}_k} \left(\sum_{t=0}^T \left(\hat{r}_k(\underline{p}_k, \underline{C}^{int}(t_i)) - r_k(t_i) \right)^2 \right) \quad (\text{формула 27})$$

Согласованные таким образом направления всех рассчитанных $\hat{r}_k(\underline{p}_k, \underline{C}^{int}(t))$ представлены вместе с соответствующим $r_k(t)$ на фиг. 11. Направления первых заштрихованы, направления последних представлены в виде сплошных линий. Заметно, что направление в качественном отношении совпадает. Это значит, что с помощью выбранных кинетик также можно хорошо показать динамику процесса. Эта информация на этом этапе моделирования очень необходима, так как при неудовлетворительном воспроизведении можно повторить быстро проводимые этапы g) (выбор других кинетик) и h) (оценка параметров), пока не достигнут необходимой степени согласования. Здесь также не требовался этап i).

9. Этап j).

Последующее согласование значений параметров модели $\underline{p}$ проводят с помощью данных измерения из этапа c). Для этого оптимизируют одновременно все параметры. Кроме того, включают не рассматриваемые ранее процессы апоптоза и лизиса. Они требуются для дифференциальных уравнений, которые описывают развитие количества жизнеспособных клеток и общее количество клеток:

$$\frac{dX_v}{dt} = (\mu_x - \mu_d) \cdot X_v \quad (\text{формула 28})$$

$$\frac{dX_t}{dt} = \mu_x \cdot X_v - K_l \cdot (X_t - X_v) \quad (\text{формула 29})$$

Выбранная кинетика для описания апоптоза:

$$\mu_d(t) = \mu_{d,max} \cdot \frac{([Lac](t) - C_{Lac,cr})}{K_{d,Lac} + ([Lac](t) - C_{Lac,cr})}, [Lac] \geq C_{Lac,cr} \quad (\text{формула 30})$$

$$\mu_d(t) = 0, [Lac] < C_{Lac,cr} \quad (\text{формула 31})$$

Норма лизиса K_l является в процессе постоянной. Кроме параметров норм реакции на этом этапе определяют введенные через апоптоз и лизис параметры $C_{Lac,cr}$ (критическую концентрацию лактата), $\mu_{d,max}$ (максимальный показатель гибели микроорганизмов), $K_{d,Lac}$ (параметр Моно для описания влияния концентрации лактата на показатель гибели микроорганизмов) и K_l (норму лизиса). В примере, исходя из начальных параметров набора данных с помощью цифрового решения ODE-системы, устанавливают направление оцениваемой концентрации $\hat{C}(t)$. При этом свели к минимуму разницу между определенными концентрациями $C^m(t)$, оцениваемыми концентрациями $\hat{C}(t)$ с помощью обычных методов со следующей целевой функцией:

$$\min_{\underline{p}} \left(\sum_{i=1}^{n_{comp}} \left(\sum_{l=0}^T \left(\hat{C}_i(p, t_l) - C_i^m(t_l) \right)^2 \right) \right) \quad (\text{формула 32})$$

Имея 33 параметра $\underline{p}$, эту оптимизацию, как правило, провести сложно, так как целевая функция имеет много локальных наилучших вариантов (оптимумов). Если запускают детерминированный алгоритм оптимизации, как, например, алгоритм Левенберга-Марквардта с известными из этапа h) начальными значениями параметров, то шансы на успех сильно повышаются. Согласованный ход технологического процесса представлен на фиг. 12. Согласованные параметры представлены в табл. 5.

Таблица 5. Параметры кинетик, а также апоптоза и лизиса

$K_{m,Glc,1}$	14,6	$K_{m,Gln,7}$	0,0187
$K_{m,Ala,1}$	3,41	$K_{m,Asp,7}$	0,872
$K_{m,Glc,2}$	0,0508	$r_{1,max}$	9,47
$K_{m,Gln,2}$	0,00881	$r_{2,max}$	9,91
$K_{m,Asn,2}$	1,38	$r_{3,max}$	57,6
$K_{m,Ala,2}$	2,19	$r_{4,max}$	21,7
$K_{m,Glc,3}$	7,13	$r_{5,max}$	0,345
$K_{m,Asn,3}$	6,84	$r_{6,max}$	49,4
$K_{m,Xt,4}$	0,0315	$r_{7,max}$	3,03
$K_{m,Glc,4}$	1,29	$K_{l,Asn,1}$	16,1
$K_{m,Gln,4}$	2,19	$K_{l,Lac,3}$	0,681
$K_{m,Asn,4}$	1,68	$K_{l,Asp,5}$	9,74
$K_{m,Glc,5}$	100	$K_{l,Asn,5}$	1,10
$K_{m,Gln,5}$	28,2	$\mu_{d,max}$	0,125
$K_{m,Asp,5}$	102	$K_{d,Lac}$	1,01
$K_{m,Glc,6}$	0,0451	$C_{Lac,cr}$	1,22
$K_{m,Gln,6}$	0,791	K_l	0,00843
$K_{m,Asp,6}$	1,06	$K_{m,Glc,7}$	0,0145

10. Этап k).

Результатом является модель, состоящая из матрицы L, кинетик из табл. 4, а также кинетик апоптоза с принадлежащими к ним значениями параметров из табл. 5.

Перечень символов.

<u>—</u> (подчеркивание)	обозначает вектор
i (индекс i)	обозначает i -элемент вектора
k (индекс k)	обозначает k -элемент вектора
[]	обозначает концентрацию стоящего в скобках компонента
C	концентрация
ΔC	разница концентраций
C^{Int}	интерполированная концентрация
$\hat{C}$	оцениваемая концентрация (например, с помощью решения дифференциального уравнения)
C_s	концентрация со сдвигом
C_{cr}	критическая концентрация
C^m	определенная концентрация
D	степень разбавления
q	Установленные специфические для клеток нормы выхода и поступления
$\tilde{q}$	установленные специфические для клеток нормы выхода и поступления, которые рассчитали от любой единицы как $\frac{[Stoffmenge]}{[Zeit-Zellzahl]}$
r	установленная скорость реакции
$\hat{r}$	оцениваемая скорость реакции (например, с помощью вычисления кинетики реакции)
$\tilde{r}$	ограничение кинетики
r_{max}	параметр кинетики химических реакций
N	стехиометрическая матрица
N_p	внешняя стехиометрическая матрица
K	матрица, которая содержит макрореакции
E	матрица, которая содержит все элементарные волны
X_t	общее число клеток
X_v	число живых клеток
μ	скорости роста
μ_d	показатель гибели микроорганизмов
$\tilde{\mu}$	скорость роста, рассчитанная от любой единицы до $\frac{[Stoffmenge]}{[Zeit-Zellzahl]}$
K_d	скорость лизиса

K_I	параметр ингибирования-лимитирования
K_M	параметр лимитирования субстрата
n	параметр Хилла лимитирования ингибирования или субстрата
L	подмножество макрореакций, которое используют для модели
p	параметры модели
S	субстрат
SSR_q	сумма возведенных в квадрат остатков особых норм выхода и поступления
SSR_c	сумма возведенных в квадрат остатков концентраций
SSR_r	сумма возведенных в квадрат остатков скоростей реакций

Описание фигур

Фиг. 1 показывает сдвиг данных измерений. Представлено фактическое направление измеряемой величины ($C_i(t)$), которое резко изменяется при изменениях скоростей разбавления ($D(t)$). Направление со сдвигом ($C_{i,s}(t)$) отображает только вызванные клеткой изменения.

Фиг. 2 показывает Flux Map двух специфических норм q_1 и q_2 . Горизонталы показывают частоту, с которой случается соответствующая комбинация норм в определенных данных.

Фиг. 3 показывает трехмерное изображение пространства решений, которое включает положительную линейную комбинацию. Черным изображено пространство решений всего комплекта, серым - одно-го подмножества.

Фиг. 4 показывает Flux Map двух специфических норм q_1 и q_2 . В виде векторов представлены 2-мерные проекции макрореакций одного комплекта L .

Фиг. 5 показывает схематичное изображение метаболической сети из Niu и др. При этом ограничение клеток показано в виде корова. Внутриклеточное разграничение митохондрии показано заштрихованной линией. Внешние компоненты обозначают индексом "xt". Стрелки и стрелки, обозначенные точками, означают реакции.

Фиг. 6 показывает данные измерений ферментации с клетками гибридомы из Vaughman и др. Общее число клеток (все клетки) при этом рассчитывают из суммы живых клеток (vital Cells) и мертвых клеток (dead Cells). Сокращения GLC, GLN, ASP, ASN, LAC, ALA и PRO означают субстрат глюкозы и аминокислот глутамина, аспаргиновой кислоты, аспарагина, аланина и пролина, а также продукта обмена веществ лактата. Сокращение MAB означает продукт моноклональных антител и BM d-биомассу.

Фиг. 7 показывает скорости роста и показатель гибели микроорганизмов, а также особые для клеток скорости поступления и выхода. Все особые для клеток скорости до q_{MAB} указаны в $\left[\frac{mM}{h \cdot 10^9 Cells} \right]$. Скорость q_{MAB} указана в $\left[\frac{10^{-4} mM}{h \cdot 10^9 Cells} \right]$.

Фиг. 8 показывает приближенные к "линейной оценке скоростей реакции выбранных макрореакций" концентрации с выбранным набором реакций. Для этого общее число клеток (X_t), а также концентрацию антител (MAB) пересчитывают на C-моль.

Фиг. 9 показывает самую маленькую установленную сумму квадратичной погрешности ("Minimum error"), нанесенную над числом макрореакций в подмножестве (n_R).

Фиг. 10 показывает установленные с помощью метода согласно изобретению "линейного оценивания скоростей реакций выбранных макрореакций" скорости макрореакций $r(t)$.

Фиг. 11 показывает установленные помощью метода согласно изобретению "линейного оценивания скоростей реакций выбранных макрореакций" скорости макрореакций $r(t)$ (сплошная линия) вместе с алгебраически рассчитанными скоростями реакции $\hat{r}(p, \underline{C}^{int}(t))$ (пунктирная линия).

Фиг. 12 показывает сравнение определенных концентраций $\underline{C}^m(t)$ (точки) и имитированного направления процесса $\hat{\underline{C}}(t)$ (сплошная линия). Концентрации указаны в [mM]. Исключениями являются количество жизнеспособных клеток и общее количество клеток (X_v/X_t в $[10^9 \text{ cells/l}]$) и концентрация антител (mAb в $[10^{-4} \text{ mM}]$).

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Компьютеризованный способ для контроля и проведения технологического процесса биореакции с организмом, включающий следующие этапы:

создание модели упомянутой биореакции с организмом, состоящей из матрицы L , кинетик $\hat{r}_k$ и значений параметров модели $\hat{p}$, включающее следующие этапы:

а) выбор путей обмена веществ организма, причем пути обмена веществ организма представляют собой внутриклеточные реакции, состоящие из субстратов, метаболитов, продуктов и биохимических реакций между ними, ввод их стехиометрических свойств и их свойств обратимости реакций в способ в виде базовых сведений и расчет элементарных мод (EMs) на основе этих вводимых данных, причем каждая элементарная мода (ЕМ) является линейной комбинацией скоростей реакций из путей обмена веществ, в которой как выполняется условие установившегося режима для внутренних метаболитов, так и учитывается обратимость или необратимость реакций;

б) элементарные моды из этапа а) объединяют в матрице K , причем элементарные моды объединяют пути обмена веществ из этапа а) в макрореакциях и матрица K содержит стехиометрические свойства, а также свойства обратимости реакций всех макрореакций;

с) данные измерений биореакции с организмом, которые происходят из сенсорного контроля биореакции и анализа образцов из биореакции, содержащие данные о числе жизнеспособных клеток, об общем числе клеток и концентрациях среды, вводят;

д) число жизнеспособных клеток, общее число клеток и концентрации среды из этапа с) интерполируют с помощью метода интерполяции и из интерполированных значений рассчитывают специфические для организма скорости - скорость выхода и скорость поступления одного или более входных параметров и выходных параметров $q(t)$ для введенных путей обмена веществ из этапа а) из векторных значений $q_i(t)$ компоненты i согласно следующей формуле

$$\frac{dC_i}{dt}(t) = D(t) \cdot (C_{i,in} - C_i(t)) + q_i(t) \cdot X_v(t)$$

причем

$D(t)$ означает скорость разбавления, $C_{i,in}$ означает концентрацию компонента i при подаче, $X_v(t)$ означает число жизнеспособных клеток, $C_i(t)$ означает концентрацию компонента, $\frac{dC_i}{dt}(t)$ означает протекание временного изменения $C_i(t)$, и $\frac{dC_i}{dt}(t)$ определяют из интерполированных сплайнами значений $C_i(t)$;

е) соответствующие макрореакции выбирают в виде подмножества элементарных мод из этапа а) посредством

i) независимого от данных предварительного сокращения количества элементарных мод из этапа а) путем геометрической редукции, причем рассчитывают повторно для случайно выбранной элементарной моды (ЕМ) все косинус-подобия ко всем другим модам, а ЕМ с наибольшим подобием удаляют из подмножества, пока не получают предварительно определенное количество элементарных мод (EMs), и/или зависимо от данных предварительного сокращения путем сравнения коэффициентов выхода продукции EMs (Y^{EM}) с коэффициентами выхода продукции, которые рассчитывают из специфических скоростей $q(t)$ из этапа д) (Y^m),

ii) выбора подмножества из предварительного сокращения из этапа е) i) с данными измерений из этапа с) и/или одной или более скоростями из этапа д) с помощью алгоритма согласно математическому критерию качества и объединения подмножества в матрице L , причем в качестве алгоритма используют либо набор с оптимальным значением критерия качества случайно выбранных EMs, либо эволюционированный алгоритм,

iii) при необходимости подмножества отображают графически;

ф) с помощью метода интерполяции рассчитывают скорости реакции макрореакции подмножества $r(t)$ на основе предоставленных данных измерений из этапа с) с помощью линейной оценки объемных скоростей реакции через промежуток времени и/или из скоростей реакций из этапа д) с помощью минимизирования отклонения скорости реакции из этапа д) для каждого рассматриваемого момента времени t_i с помощью алгоритма неотрицательных квадратов (NNLS)

$$\min_{r(t_i)} (\|L \cdot r(t_i) - \hat{q}(t_i)\|)$$

г) кинетики макрореакций подмножества из этапа е) ii) проектируют на следующих промежуточных этапах; благодаря этому определяют параметры модели,

и) из стехиометрии макрореакций проектируют родовые кинетики $\hat{r}_k$ посредством того, что для всех субстратов i всех макрореакций k предполагают лимитирование $\hat{r}_i(t)$ по формуле

$$\hat{r}_k(t) = r_{k,max} \cdot \prod_{i=1}^{N_t} \hat{r}_i(t) = r_{k,max} \prod_{i=1}^{N_t} \left(\frac{C_i(t)}{K_{m,k,i} + C_i(t)} \right)^{n_i} \quad (a)$$

- причем постоянные Монода $K_{m,k,i}$ и коэффициенты Хилла n_i представляют параметры уравнения,
- ii) параметры воздействия на макрореакции определяют из скоростей реакций из этапа f) с помощью определения корреляции между состоянием процесса и скоростями реакции $r(t)$ из этапа d) посредством статистического метода,
- iii) родовые кинетики из этапа g) i) пополняются элементами, которые дают количественную оценку установленным на этапе g) ii) параметрам воздействия;
- h) при необходимости, для кинетик $\hat{r}_k$ из этапа g) проводят первое согласование значений параметров модели $\underline{p}$ отдельно для каждой макрореакции с рассчитанными скоростями реакций $r(t)$ из этапа f) отдельно для каждой макрореакции:

$$\min_{\underline{p}} \sum_{k=1}^{N_t} (\hat{r}_k(\underline{p}) - r_k)^2$$

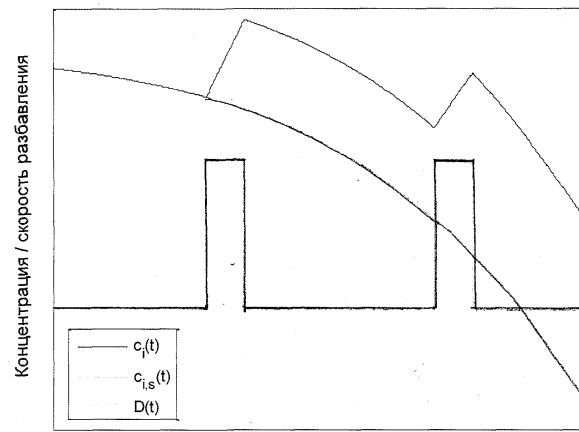
- i) при необходимости повторяют этапы g) и h) до получения заданного согласования;
- j) значения параметров модели $\underline{p}$ согласуют с данными измерений из этапа c);
- k) матрица L , кинетики из этапа g) и значения параметров модели из этапа j) образуют модель и выдаются, передаются в методы оценки состояния, основанной на модели, в модуль управления технологическим процессом или разработки технологического процесса и/или используются в основанном на модели предикативном закрытом контуре регулирования для оптимального управления технологическим процессом.

2. Компьютеризованный способ по п.1, причем на этапе d) рассчитывают также скорости роста $\mu(t)$, особенно предпочтительно также показатели гибели микроорганизмов μ_d по формуле (b) и (c):

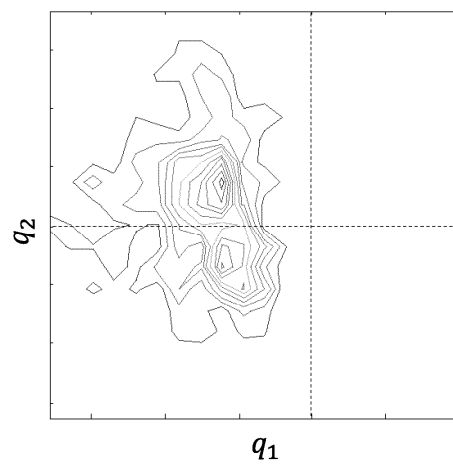
$$(b) \quad \frac{dX_t}{dt}(t) = -D(t) \cdot X_t(t) + \mu(t) \cdot X_v(t)$$

$$(c) \quad \frac{dX_v}{dt}(t) = -D(t) \cdot X_v(t) + (\mu(t) - \mu_d(t)) \cdot X_v(t).$$

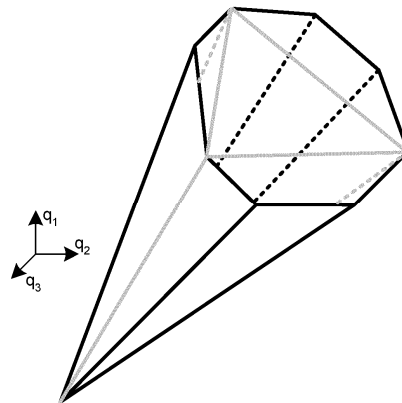
3. Компьютеризованный способ по п.1 или 2, в котором осуществляют этап h).
4. Компьютеризованный способ по одному из пп.1-3, причем на этапе e) ii) для выбора подмножества макрореакций рассчитывают критерий качества путем линейного оценивания скоростей реакций выбранных макрореакций.
5. Компьютеризованный способ по одному из пп.1-4, причем на этапе e) ii) для выбора подмножества макрореакций рассчитывают критерий качества путем линейного оценивания скоростей реакций выбранных макрореакций и используют эволюционный алгоритм для выбора.
6. Компьютеризованный способ по одному из пп.1-5, причем проводят сдвиг данных измерений перед применением метода интерполяции на этапе d) для достижения описания постоянного расхода.
7. Компьютеризованный способ по одному из пп.1-6, причем на этапе f) проводят расчет скоростей реакций $r(t)$ с помощью линейной оценки объемных скоростей реакции через промежуток времени.
8. Компьютеризованный способ по одному из пп.1-7, причем на этапе e) i) проводят зависимое от данных предварительное сокращение.
9. Компьютеризованный способ по одному из пп.1-8, причем на этапе e) ii) проводят выбор из предварительного сокращения из этапа e) i) с данными измерений из этапа c).



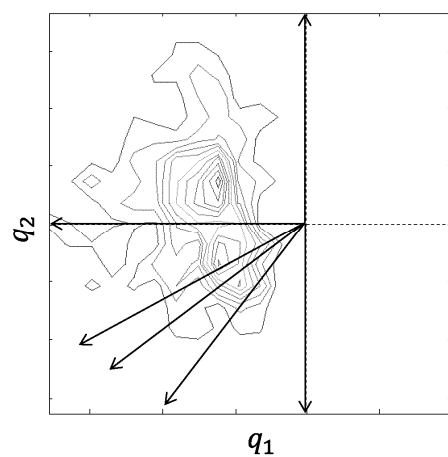
Фиг. 1



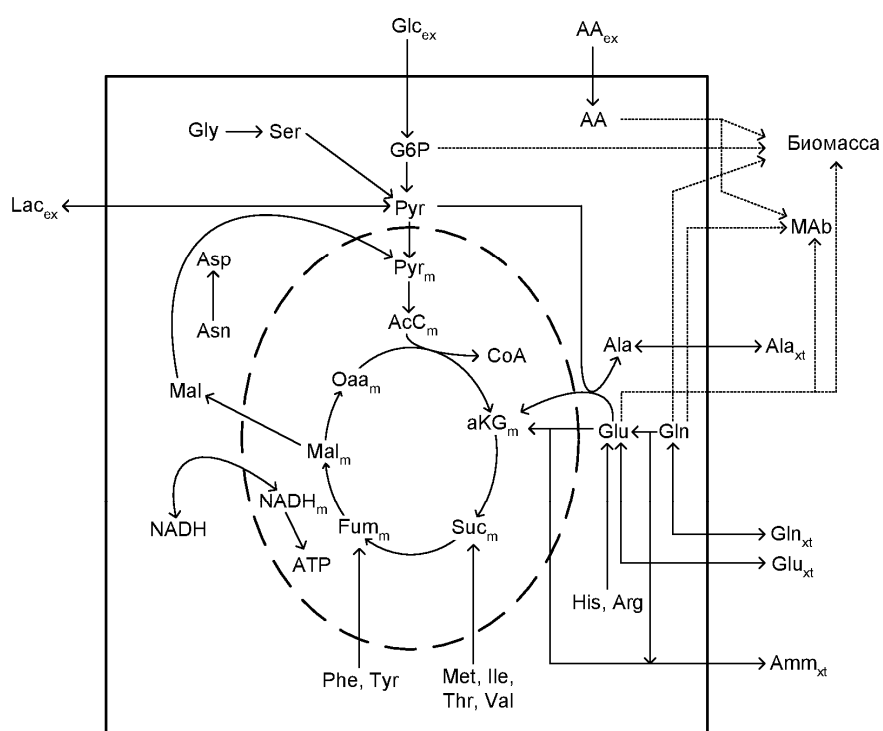
Фиг. 2



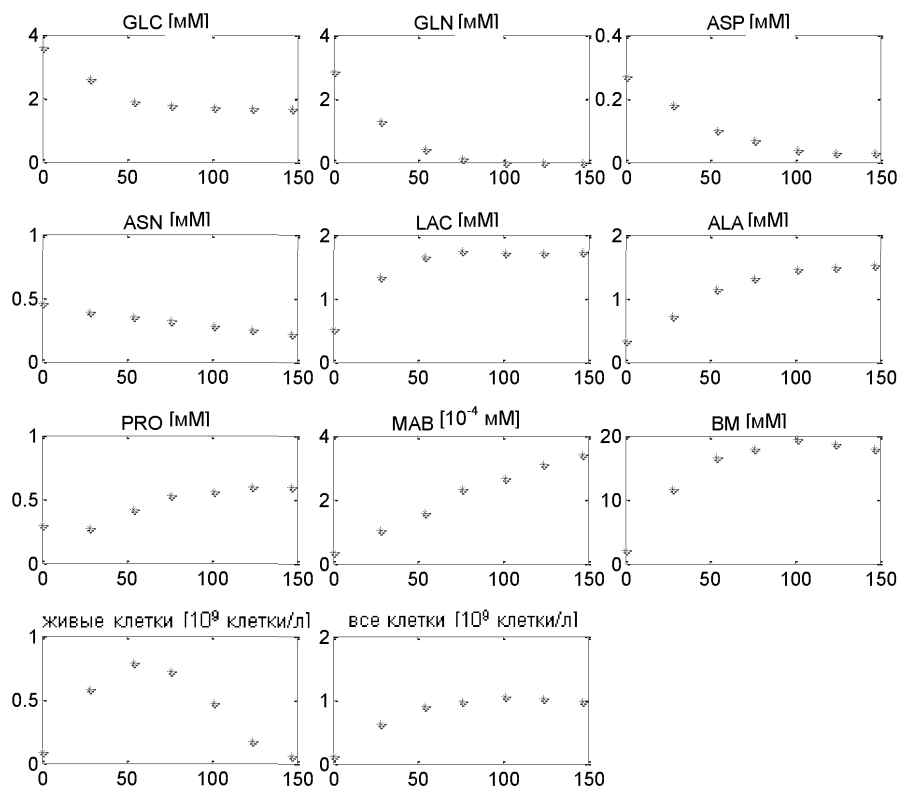
Фиг. 3



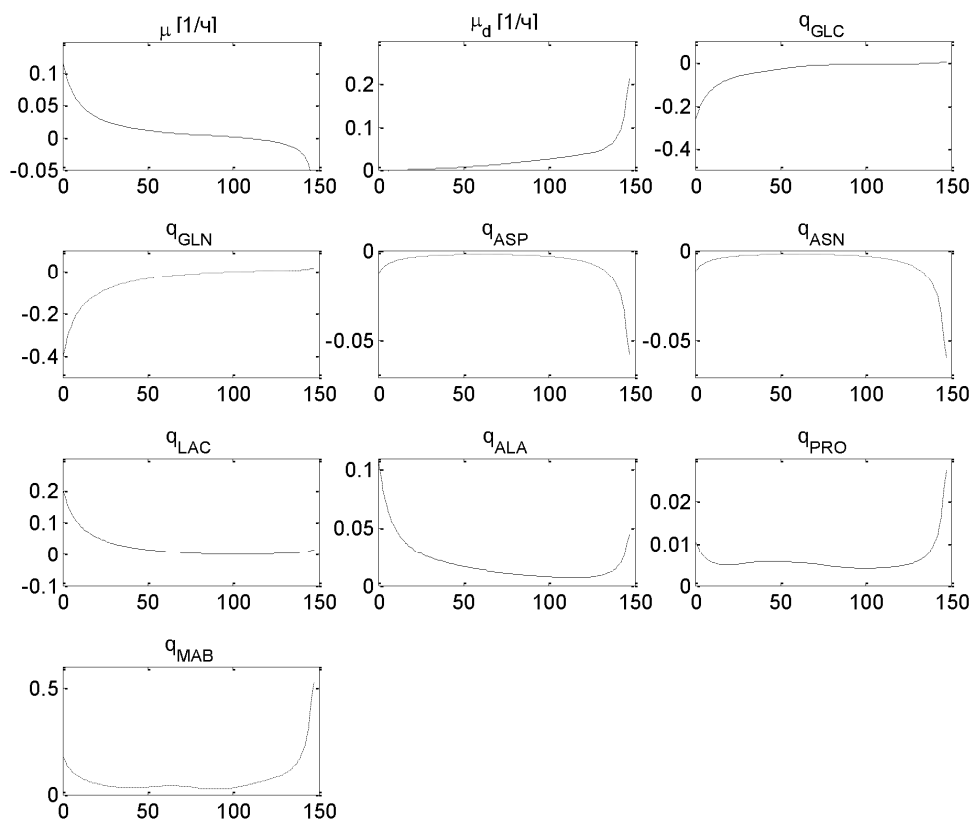
Фиг. 4



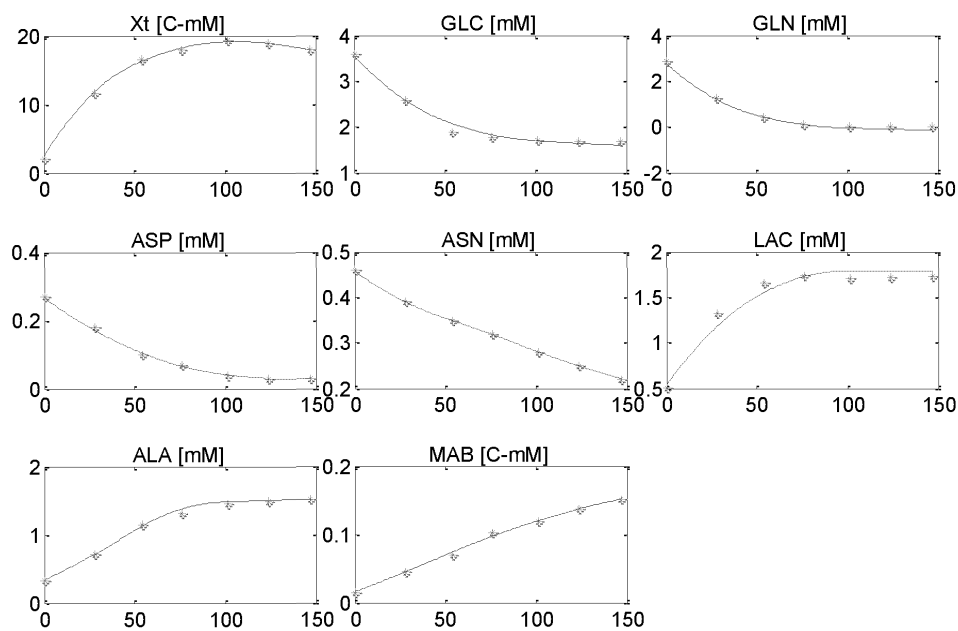
Фиг. 5



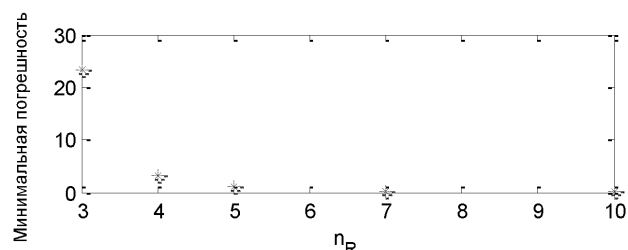
Фиг. 6



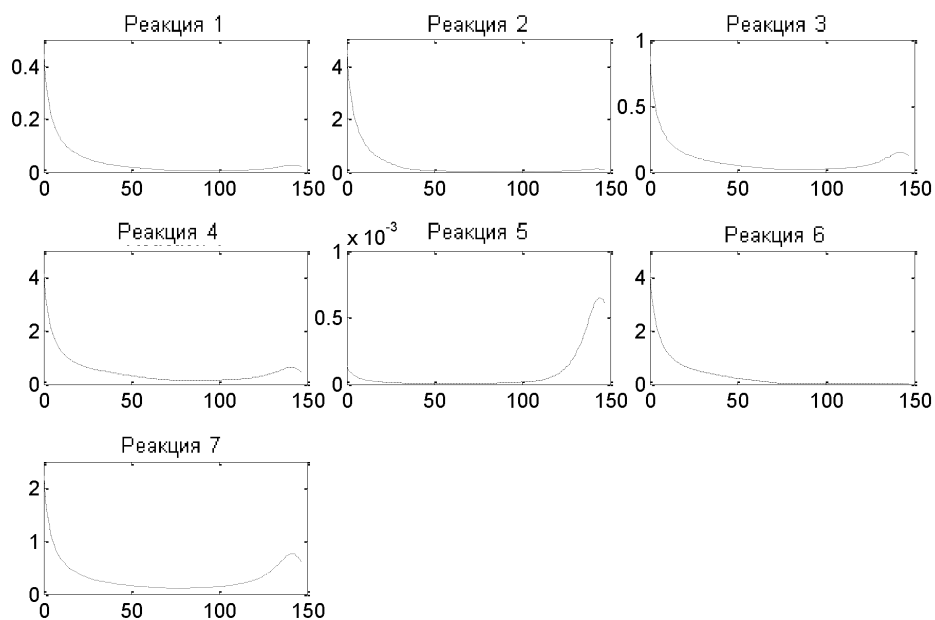
Фиг. 7



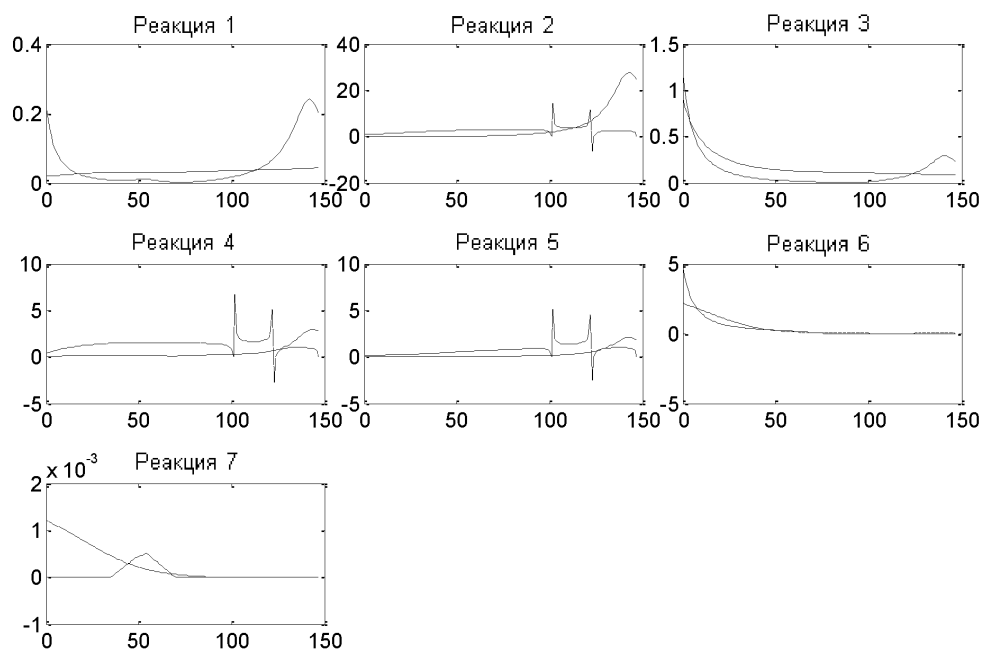
Фиг. 8



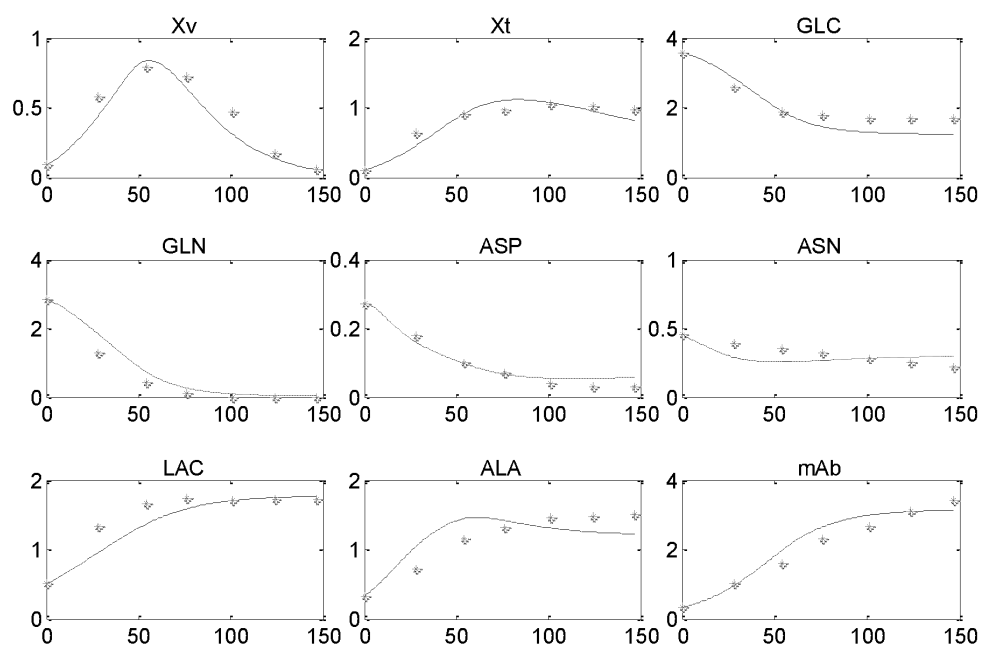
Фиг. 9



Фиг. 10



Фиг. 11



Фиг. 12



Евразийская патентная организация, ЕАПВ

Россия, 109012, Москва, Малый Черкасский пер., 2