

(19)



**Евразийское
патентное
ведомство**

(21) **202292911** (13) **A1**

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2022.12.01

(51) Int. Cl. *G01N 30/86* (2006.01)
G01N 30/88 (2006.01)

(22) Дата подачи заявки
2021.04.13

**(54) УЛЬТРАФИОЛЕТОВЫЙ КОНТРОЛЬ ХАРАКТЕРИСТИК ХРОМАТОГРАФИИ
ПОСРЕДСТВОМ МОДЕЛИ ОРТОГОНАЛЬНЫХ ЧАСТИЧНЫХ НАИМЕНЬШИХ
КВАДРАТОВ**

(31) 63/009,835

(72) Изобретатель:

(32) 2020.04.14

**Ширли Эрик, Бреннан Кевин, Мао
Натан (US)**

(33) US

(86) PCT/US2021/027043

(74) Представитель:

(87) WO 2021/211550 2021.10.21

Медведев В.Н. (RU)

(71) Заявитель:

**РИДЖЕНЕРОН
ФАРМАСЬЮТИКАЛЗ, ИНК. (US)**

(57) Описаны способы контроля характеристик колоночной хроматографии. Способы могут включать в себя получение одной или более хроматограмм в ультрафиолетовом спектре (УФ-хроматограмм), сгенерированных системой хроматографии во время очистки и/или разделения образца; и анализ одной или более полученных УФ-хроматограмм с помощью модели ортогональных частичных наименьших квадратов (OPLS), таким образом обеспечивая обнаружение ухудшения состояния колонки до разрушения колонки и количественный анализ УФ-сигнала в указанных одной или более УФ-хроматограмм.



A1

202292911

202292911

A1

ОПИСАНИЕ ИЗОБРЕТЕНИЯ

2420-576025EA/032

УЛЬТРАФИОЛЕТОВЫЙ КОНТРОЛЬ ХАРАКТЕРИСТИК ХРОМАТОГРАФИИ ПОСРЕДСТВОМ МОДЕЛИ ОРТОГОНАЛЬНЫХ ЧАСТИЧНЫХ НАИМЕНЬШИХ КВАДРАТОВ

ОБЛАСТЬ ТЕХНИКИ

[001] Настоящее изобретение относится к хроматографии и, в частности, к способу ультрафиолетового (УФ) контроля характеристик хроматографии, например, путем применения модели ортогональных частичных наименьших квадратов (OPLS; orthogonal partial least squares).

УРОВЕНЬ ТЕХНИКИ

[002] Хроматография представляет собой метод отделения, применяемый для очистки белков, включая антитела, в биофармацевтической промышленности. В хроматографии применяют колонку, упакованную небольшими частицами (смолой), которая взаимодействует с продуктом (требуемым белком) для отделения его от примесей. Хроматограммы представляют собой графические записи хроматографического разделения для обнаружения того, как такие параметры, как УФ поглощение, проводимость и рН, изменяются с течением времени во время прогона хроматографии. Состояние упакованного слоя смолы в хроматографической колонке может ухудшаться с течением времени, что может оказывать воздействие на эффективность хроматографического разделения и влиять на качество продукта.

[003] Полезно знать, когда состояние хроматографической колонки начинает ухудшаться, чтобы ее можно было переупаковать до того, как состояние колонки ухудшится до того момента, когда это может влиять на качество продукта. Признаки ухудшения состояния колонки часто проявляются в УФ поглощении в хроматограмме; однако признаки ухудшения состояния могут быть сложно визуально идентифицировать. Анализ белков на основе хроматографии может значительно выиграть от усовершенствованных способов обнаружения и контроля ухудшения состояния колонки, что в результате может привести к улучшенной очистке белков, включая улучшенную очистку антител.

КРАТКОЕ ОПИСАНИЕ ИЗОБРЕТЕНИЯ

[004] В одном аспекте настоящее изобретение обеспечивает способ контроля характеристик колоночной хроматографии, включающий: получение одной или более хроматограмм в ультрафиолетовом спектре (УФ-хроматограмм), генерируемых системой хроматографии во время очистки и/или разделения образца; и анализ одной или более полученных УФ-хроматограмм с помощью модели ортогональных частичных наименьших квадратов (OPLS), таким образом обеспечивая обнаружение ухудшения состояния колонки до разрушения колонки и количественный анализ УФ-сигнала в указанных одной или более УФ-хроматограмм.

[005] В некоторых вариантах осуществления способ дополнительно включает в

себя создание модели OPLS.

[006] В некоторых вариантах осуществления создание модели OPLS включает в себя: выбор процесса или однократной операции модели OPLS; сбор исходных данных для УФ-кривой одного или более колоночных циклов/партий/прогонов для выбранного процесса или однократной операции; нормализацию и совмещение собранных исходных данных; необязательно, генерирование искусственно созданных кривых на основании нормализованных исходных данных; классификацию и форматирование данных для импортирования во многомерный инструмент; импортирование классифицированных и отформатированных данных во многомерный инструмент для генерирования обучающего набора; и генерирование модели OPLS.

[007] В некоторых вариантах осуществления способ дополнительно включает в себя оптимизацию сгенерированной модели OPLS, подтверждение и тестирование оптимизированной модели OPLS.

[008] В некоторых вариантах осуществления процесс или однократная операция представляет собой однократную операцию хроматографии для одной молекулы.

[009] В некоторых вариантах осуществления процесс или однократная операция представляет собой стадию аффинной хроматографии белков для одной молекулы.

[0010] В некоторых вариантах осуществления сбор исходных данных для УФ-кривой одного или более колоночных циклов/партий/прогонов для выбранного процесса или однократной операции включает в себя сбор значений УФ-поглощения в соответствующих объемах колонки для одного или более циклов/партий/прогонов.

[0011] В некоторых вариантах осуществления нормализация и совмещение собранных исходных данных включает в себя нормализацию и совмещение значений УФ-поглощения и совмещение объемов колонки.

[0012] В некоторых вариантах осуществления нормализованные значения УФ-поглощения включают в себя удаление вариации различий по величине в необработанном УФ-сигнале.

[0013] В некоторых вариантах осуществления, необязательно, генерирование искусственно созданных кривых на основании нормализованных исходных данных включает в себя генерирование искусственно созданных кривых на основании нормализованных исходных данных, когда одна или более неприемлемых УФ-хроматограмм не доступны.

[0014] В некоторых вариантах осуществления способ дополнительно включает в себя обеспечение образца в системе хроматографии перед получением одной или более хроматограмм в ультрафиолетовом спектре (УФ-хроматограмм), сгенерированных системой хроматографии во время очистки и/или разделения образца.

[0015] В некоторых вариантах осуществления система хроматографии представляет собой систему жидкостной хроматографии.

[0016] В некоторых вариантах осуществления образец содержит белок.

[0017] В некоторых вариантах осуществления белок представляет собой антитело,

слитый белок, рекомбинантный белок или их комбинацию.

[0018] В некоторых вариантах осуществления антитело представляет собой моноклональное антитело.

[0019] В некоторых вариантах осуществления моноклональное антитело имеет изотип IgG1, IgG2, IgG3, IgG4 или смешанный изотип.

[0020] Также описан способ создания модели ортогональных частичных наименьших квадратов (OPLS) для УФ-контроля характеристик хроматографической колонки, включающий в себя: выбор процесса или однократной операции модели OPLS; сбор исходных данных для хроматограммы в ультрафиолетовом спектре (УФ-хроматограммы) одного или более колоночных циклов/партий/прогонов для выбранного процесса или однократной операции; нормализацию и совмещение собранных исходных данных; необязательно, генерирование искусственно созданных кривых на основании нормализованных исходных данных; классификацию и форматирование данных для импортирования во многомерный инструмент; импортирование классифицированных и отформатированных данных во многомерный инструмент для генерирования обучающего набора; и генерирование модели OPLS.

[0021] В некоторых вариантах осуществления способ дополнительно включает в себя оптимизацию сгенерированной модели OPLS, подтверждение и тестирование оптимизированной модели OPLS.

[0022] В некоторых вариантах осуществления процесс или однократная операция представляет собой однократную операцию хроматографии для одной молекулы.

[0023] В некоторых вариантах осуществления процесс или однократная операция представляет собой стадию аффинной хроматографии белков для одной молекулы.

[0024] В некоторых вариантах осуществления сбор исходных данных для УФ-кривой одного или более колоночных циклов/партий/прогонов для выбранного процесса или однократной операции включает в себя сбор значений УФ-поглощения в соответствующих объемах колонки для одного или более циклов/партий/прогонов.

[0025] В некоторых вариантах осуществления нормализация и совмещение собранных исходных данных включает в себя нормализацию и совмещение значений УФ-поглощения и совмещение объемов колонки.

[0026] В некоторых вариантах осуществления нормализованные значения УФ-поглощения включают в себя удаление вариации различий по величине в необработанном УФ-сигнале.

[0027] В некоторых вариантах осуществления, необязательно, генерирование искусственно созданных кривых на основании нормализованных исходных данных включает в себя генерирование искусственно созданных кривых на основании нормализованных исходных данных, когда одна или более неприемлемых УФ-хроматограмм не доступны.

[0028] В некоторых вариантах осуществления способ дополнительно включает в себя обеспечение образца в системе хроматографии перед получением одной или более

хроматограмм в ультрафиолетовом спектре (УФ-хроматограмм), сгенерированных системой хроматографии во время очистки и/или разделения образца.

[0029] В некоторых вариантах осуществления система хроматографии представляет собой систему жидкостной хроматографии.

[0030] В некоторых вариантах осуществления образец содержит белок.

[0031] В некоторых вариантах осуществления белок представляет собой антитело, слитый белок, рекомбинантный белок или их комбинацию.

[0032] В некоторых вариантах осуществления антитело представляет собой моноклональное антитело.

[0033] В некоторых вариантах осуществления моноклональное антитело имеет изотип IgG1, IgG2, IgG3, IgG4 или смешанный изотип.

[0034] В вариантах осуществления обеспечен энергонезависимый машиночитаемый носитель данных с хранимой на нем исполняемой программой для контроля характеристик колоночной хроматографии, причем программа предписывает микропроцессору выполнять следующие стадии: получение одной или более хроматограмм в ультрафиолетовом спектре (УФ-хроматограмм), сгенерированных системой хроматографии во время очистки и/или разделения образца; и анализ одной или более полученных УФ-хроматограмм с помощью модели ортогональных частичных наименьших квадратов (OPLS), таким образом обеспечивая обнаружение ухудшения состояния колонки до разрушения колонки и количественный анализ УФ-сигнала в указанных одной или более УФ-хроматограмм.

[0035] В некоторых вариантах осуществления энергонезависимый машиночитаемый носитель данных дополнительно содержит команды по созданию модели OPLS.

[0036] В некоторых вариантах осуществления создание модели OPLS включает в себя: выбор процесса или однократной операции модели OPLS; сбор исходных данных для УФ-кривой одного или более колоночных циклов/партий/прогонов для выбранного процесса или однократной операции; нормализацию и совмещение собранных исходных данных; необязательно, генерирование искусственно созданных кривых на основании нормализованных исходных данных; классификацию и форматирование данных для импортирования во многомерный инструмент; импортирование классифицированных и отформатированных данных во многомерный инструмент для генерирования обучающего набора; и генерирование модели OPLS.

[0037] В некоторых вариантах осуществления способ дополнительно включает в себя оптимизацию сгенерированной модели OPLS, подтверждение и тестирование оптимизированной модели OPLS.

[0038] В некоторых вариантах осуществления процесс или однократная операция представляет собой однократную операцию хроматографии для одной молекулы.

[0039] В некоторых вариантах осуществления процесс или однократная операция представляет собой стадию аффинной хроматографии белков для одной молекулы.

[0040] В некоторых вариантах осуществления сбор исходных данных для УФ-кривой одного или более колоночных циклов/партий/прогонов для выбранного процесса или однократной операции включает в себя сбор значений УФ-поглощения в соответствующих объемах колонки для одного или более циклов/партий/прогонов.

[0041] В некоторых вариантах осуществления нормализация и совмещение собранных исходных данных включает в себя нормализацию и совмещение значений УФ-поглощения и совмещение объемов колонки.

[0042] В некоторых вариантах осуществления нормализованные значения УФ-поглощения включают в себя удаление вариации различий по величине в необработанном УФ-сигнале.

[0043] В некоторых вариантах осуществления, необязательно, генерирование искусственно созданных кривых на основании нормализованных исходных данных включает в себя генерирование искусственно созданных кривых на основании нормализованных исходных данных, когда одна или более неприемлемых УФ-хроматограмм не доступны.

[0044] В некоторых вариантах осуществления способ дополнительно включает в себя обеспечение образца в системе хроматографии перед получением одной или более хроматограмм в ультрафиолетовом спектре (УФ-хроматограмм), сгенерированных системой хроматографии во время очистки и/или разделения образца.

[0045] В некоторых вариантах осуществления система хроматографии представляет собой систему жидкостной хроматографии.

[0046] В некоторых вариантах осуществления образец содержит белок.

[0047] В некоторых вариантах осуществления белок представляет собой антитело, слитый белок, рекомбинантный белок или их комбинацию.

[0048] В некоторых вариантах осуществления антитело представляет собой моноклональное антитело.

[0049] В некоторых вариантах осуществления моноклональное антитело имеет изотип IgG1, IgG2, IgG3, IgG4 или смешанный изотип.

[0050] В различных вариантах реализации любые из признаков или компонентов вариантов реализации, обсуждаемых выше или в настоящем документе, можно комбинировать, и такие комбинации входят в рамки настоящего изобретения. Любое конкретное значение, обсуждаемое выше или далее в настоящем документе, может быть комбинировано с другим взаимосвязанным значением, обсуждаемым выше или далее в настоящем документе, для определения диапазона, в котором указанные значения представляют собой верхний и нижний пределы диапазона, и такие диапазоны и все значения, входящие в такие диапазоны, включены в объем настоящего изобретения. Каждое из значений, обсуждаемых выше или далее в настоящем документе, может быть указано с отклонением в 1%, 5%, 10% или 20%. Другие варианты реализации будут очевидны из обзора следующего подробного описания изобретения.

КРАТКОЕ ОПИСАНИЕ ГРАФИЧЕСКИХ МАТЕРИАЛОВ

[0051] На **Фиг. 1А** показан иллюстративный рабочий процесс в соответствии с вариантами осуществления, описанными в настоящем документе.

[0052] На **Фиг. 1В** показана геометрическая интерпретация OPLS.

[0053] На **Фиг. 1С** показано схематическое изображение, иллюстрирующее прогноз OPLS.

[0054] На **Фиг. 1D** показана блок-схема, иллюстрирующая иллюстративный способ создания УФ-модели OPLS.

[0055] На **Фиг. 2** показана иллюстративная хроматограмма связывания с элюированием с выделенным блоком элюирования.

[0056] На **Фиг. 3** показана иллюстративная проточная хроматограмма с выделенным УФ-отрывом.

[0057] На **Фиг. 4** показаны иллюстративные экспортированные исходные данные, сгенерированные в соответствии с вариантами осуществления, описанными в настоящем документе.

[0058] На **Фиг. 5-1 и Фиг. 5-2** показана иллюстративная нормализация данных в соответствии с вариантами осуществления, описанными в настоящем документе.

[0059] На **Фиг. 6-1, 6-2, 6-3, 7, 8, 9-1, 9-2, 10-1 и 10-2** показано создание искусственно сгенерированных кривых на основании нормализованных данных в соответствии с вариантами осуществления, описанными в настоящем документе.

[0060] На **Фиг. 11-1 и Фиг. 11-2** показана иллюстративная классификация и форматирование данных для импортирования Simca в соответствии с вариантами осуществления, описанными в настоящем документе.

[0061] На **Фиг. 12А, 12В, 13А, 13В-1, 13В-2, 14А, 14В, 15А, 15В-1, 15В-2 и 16** показано иллюстративное импортирование данных в многомерный инструмент в соответствии с вариантами осуществления, описанными в настоящем документе.

[0062] На **Фиг. 17, 18, 19 и 20** показано иллюстративное генерирование модели OPLS в соответствии с вариантами осуществления, описанными в настоящем документе.

[0063] На **Фиг. 21, 22А, 22В, 23А, 23В, 24А, 24В, 25-1 и 25-2** показана иллюстративная оптимизация модели OPLS в соответствии с вариантами осуществления, описанными в настоящем документе.

[0064] На **Фиг. 26А, 26В, 27А, 27В, 28, 29-1, 29-2, 30, 31-1 и 31-2** показано иллюстративное применение созданной модели OPLS для классификации новых хроматограмм в качестве приемлемых или неприемлемых в соответствии с вариантами осуществления, описанными в настоящем документе.

[0065] На **Фиг. 32-1, 32-2, 33, 34, 35, 36, 37, 38А, 38В, 39А, 39В и 40** показаны результаты, сгенерированные с применением моделей OPLS в соответствии с вариантами осуществления, описанными в настоящем документе.

[0066] На **Фиг. 41** показано схематическое изображение иллюстративной вычислительной среды для выполнения аспектов способов, описанных в настоящем документе.

ПОДРОБНОЕ ОПИСАНИЕ ИЗОБРЕТЕНИЯ

[0067] Перед прочтением описания настоящего изобретения следует понимать, что данное изобретение не ограничивается конкретными описанными методами и экспериментальными условиями, поскольку такие методы и условия могут варьироваться. Также следует понимать, что используемая в настоящем документе терминология предназначена только для описания частных вариантов осуществления и не подразумевает ограничительного характера, так как объем настоящего изобретения ограничен только прилагаемой формулой изобретения. Любые варианты осуществления или признаки вариантов осуществления могут быть объединены друг с другом, и такие комбинации явно входят в объем настоящего изобретения.

[0068] Если не указано иное, все технические и научные термины, используемые в настоящем документе, имеют значение, обычно понимаемое специалистом в данной области техники, к которой принадлежит данное изобретение. В настоящем документе термин «около» при использовании в отношении конкретного приведенного числового значения означает, что значение может отличаться от приведенного значения не более чем на 1%. Например, в контексте настоящего документа выражение «около 100» включает 99 и 101, и все значения между ними (например, 99,1, 99,2, 99,3, 99,4 и т. д.).

[0069] Различные операции могут быть описаны как множество дискретных операций по очереди таким образом, что это может быть полезно для понимания вариантов осуществления; однако порядок описания не следует понимать как означающий, что эти операции зависят от порядка.

[0070] Могут быть использованы термины «присоединенный» и «соединенный» вместе с их производными. Эти термины не предназначены для использования как синонимы друг для друга. Скорее, в аспектах термин «присоединенный» может быть использован для указания того, что два или более элементов находятся в непосредственном физическом или электрическом контакте друг с другом. Термин «соединенный» может означать, что два или более элементов находятся в непосредственном физическом или электрическом контакте. Тем не менее, «соединенный» может также означать, что два или более элементов не находятся в непосредственном контакте друг с другом, но все еще объединяются или взаимодействуют друг с другом.

[0071] Используемые в настоящем документе термины «включают», «включает» и «включающий» означают, что они не имеют ограничительного характера, и означают «содержат», «содержит» и «содержащий» соответственно;

[0072] Хотя любые методы и материалы, подобные или эквивалентные тем, которые описаны в настоящем документе, могут использоваться при практическом осуществлении или испытании данного изобретения, предпочтительные методы и материалы описаны далее. Содержание всех патентов, заявок и непатентных публикаций, упомянутых в настоящем описании, включено в настоящий документ в полном объеме посредством ссылки.

Сокращения, используемые в настоящем документе

- [0073] СНО: Яичник китайского хомячка
- [0074] CV: Объемы колонки
- [0075] DDA: Зависимое от данных получение
- [0076] EIC: Экстрагированный ионный хроматограф
- [0077] HC Тяжелая цепь
- [0078] HIC: Хроматография гидрофобного взаимодействия
- [0079] HILIC: Жидкостная хроматография гидрофильного взаимодействия
- [0080] VM: Высокомолекулярная масса
- [0081] IgG: Иммуноглобулин G
- [0082] IPA: Изопропанол
- [0083] LC: Легкая цепь
- [0084] NM: Низкомолекулярная масса
- [0085] mAb: Моноклональное антитело
- [0086] MM: Молекулярная масса
- [0087] OPLS: Ортогональные частичные наименьшие квадраты
- [0088] ФК: Фармакокинетика
- [0089] SKO: Среднеквадратичная ошибка
- [0090] OF-ЖХ: Обращенно-фазовая жидкостная хроматография
- [0091] SME: Специалист в предметной области
- [0092] SIM: Контроль заданных ионов
- [0093] УФ: Ультрафиолетовый

Определения

[0094] В настоящем документе термин «белок» включает в себя любой аминокислотный полимер, имеющий ковалентно связанные амидные связи. Белки содержат одну или более полимерных цепей аминокислот, по существу известных в данной области как «полипептиды». «Полипептид» относится к полимеру, состоящему из аминокислотных остатков, родственных встречающихся в природе структурных вариантов и синтетических не встречающихся в природе аналогов, связанных посредством пептидных связей, их родственных встречающихся в природе структурных вариантов и синтетических не встречающихся в природе аналогов. «Синтетические пептиды или полипептиды» относятся к не встречающемуся в природе пептиду или полипептиду. Синтетические пептиды или полипептиды могут быть синтезированы, например, с применением автоматизированного синтезатора полипептидов. Специалистам в данной области известны различные способы твердофазного синтеза пептидов. Белок может содержать один или множество полипептидов с образованием одной функционирующей биомолекулы. Белок может включать в себя любой из биотерапевтических белков, рекомбинантных белков, применяемых в исследованиях или терапии, белков-ловушек и других слитых белков химерного рецептора Fc, химерных белков, антител, моноклональных антител, поликлональных антител, антител человека и

биспецифических антител. В другом иллюстративном аспекте белок может включать в себя фрагменты антител, нанотела, рекомбинантные химеры антител, цитокины, хемокины, пептидные гормоны и т. п. Белки могут быть получены с применением систем производства на основе рекомбинантных клеток, таких как баккуловирусная система насекомых, системы дрожжей (например, *Pichia sp.*), системы млекопитающих (например, клетки CHO и производные CHO, такие как клетки CHO-K1). Недавний обзор, рассматривающий биотерапевтические белки и их производство, см. в Gaderi et al., «Products platteries for biotherapeutic glycoproteins. Occurrence, impact, and challenges of non-human sialylation», (*Biotechnol. Genet. Eng. Rev.* (2012) 147-75). В некоторых вариантах осуществления белки содержат модификации, аддукты и другие ковалентно связанные фрагменты. Эти модификации, аддукты и фрагменты включают в себя, например, авидин, стрептавидин, биотин, гликаны (например, N-ацетилгалактозамин, галактозу, нейраминную кислоту, N-ацетилглюкозамин, фукозу, маннозу и другие моносахариды), ПЭГ, полигистидин, FLAGtag, мальтозасвязывающий белок (MBP), хитинсвязывающий белок (CBP), мус-эпитоп глутатион-S-трансферазы (GST), флуоресцентные метки и другие красители и т. п. Белки могут быть классифицированы на основе композиций и растворимости и, таким образом, могут включать в себя простые белки, такие как глобулярные белки и волокнистые белки; конъюгированные белки, такие как нуклеопротеины, гликопротеины, мукопротеины, хромопротеины, фосфопротеины, металлопротеины и липопротеины; и производные белки, такие как первичные производные белки и вторичные производные белки.

[0095] Вариантный белок «или вариант белка», или «вариант», используемый в данном документе, может включать в себя белок, который отличается от целевого белка в силу по меньшей мере одной модификации аминокислоты. Вариант белка может относиться к самому белку, композиции, содержащей белок, или аминокислотную последовательность, которая его кодирует. Предпочтительно, вариант белка имеет по меньшей мере одну модификацию аминокислоты по сравнению с родительским белком, например, от около одной до около десяти модификаций аминокислоты и предпочтительно от около одной до около пяти модификаций аминокислоты по сравнению с родительской. Последовательность варианта белка в настоящем документе предпочтительно будет иметь по меньшей мере около 80% гомологии с последовательностью родительского белка и наиболее предпочтительно по меньшей мере около 90% гомологии, более предпочтительно по меньшей мере около 95% гомологии. В некоторых примерах осуществления белок может представлять собой антитело, биспецифическое антитело, мультиспецифическое антитело, фрагмент антитела, моноклональное антитело или их комбинации.

[0096] Термин «антитело», используемый в настоящем документе, предназначен для обозначения молекул иммуноглобулина, состоящих из четырех полипептидных цепей, двух тяжелых (H) цепей и двух легких (L) цепей, связанных между собой дисульфидными связями (т. е., «молекул полноразмерного антитела»), а также их мультимеров (например,

IgM) или их антигенсвязывающих фрагментов. Каждая тяжелая цепь состоит из переменной области тяжелой цепи (HCVR или V_H) и константной области тяжелой цепи (состоящей из доменов C_{H1} , C_{H2} и C_{H3}). В различных вариантах осуществления тяжелая цепь может представлять собой изотип IgG. В некоторых случаях тяжелая цепь выбрана из IgG1, IgG2, IgG3 или IgG4. В некоторых вариантах осуществления тяжелая цепь имеет изотип IgG1 или IgG4, необязательно включающий в себя химерную шарнирную область изотипа IgG1/IgG2 или IgG4/IgG2. Каждая легкая цепь состоит из переменной области легкой цепи («LCVR» или « V_L ») и константной области легкой цепи (C_L). В V_H - и V_L -областях можно дополнительно выделить области гипервариабельности, называемые областями, определяющими комплементарность (CDR), которые разделены более консервативными областями, называемыми каркасными областями (FR). Каждая V_H и V_L состоит из трех CDR и четырех FR, расположенных от аминоконца к карбоксиконцу в следующем порядке: FR1, CDR1, FR2, CDR2, FR3, CDR3, FR4. Термин «антитело» включает в себя ссылку как гликозилированные, так и негликозилированные иммуноглобулины любого изотипа или подкласса. Термин «антитело» включает в себя молекулы антител, полученные, экспрессированные, созданные или выделенные рекомбинантными средствами, такими как антитела, выделенные из клетки-хозяина, трансфицированной для экспрессии антитела. Для обзора структуры антител см. Lefranc et al., *IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains*, 27(1) *Dev. Comp. Immunol.* 55-77 (2003); и M. Potter, *Structural correlates of immunoglobulin diversity*, 2(1) *Surv. Immunol. Res.* 27-42 (1983).

[0097] Термин «антитело» также включает в себя «биспецифическое антитело», который включает в себя гетеротетрамерный иммуноглобулин, который может связываться с более чем одним другим эпитопом. Одна половина биспецифического антитела, которая включает в себя одну тяжелую цепь и одну легкую цепь и шесть CDR, связывается с одним антигеном или эпитопом, а другая половина антитела связывается с другим антигеном или эпитопом. В некоторых случаях биспецифическое антитело может связываться с тем же антигеном, но в различных эпитопах или неперекрывающихся эпитопах. В некоторых случаях обе половины биспецифического антитела имеют идентичные легкие цепи, сохраняя при этом двойную специфичность. Биспецифические антитела описаны по существу в публикации патентной заявки США № 2010/0331527 (30 декабря 2010 г.).

[0098] Термин «антигенсвязывающий участок» (или «фрагмент антитела») относится к одному или более фрагментам антитела, которые сохраняют способность специфично связываться с антигеном. Примеры связывающих фрагментов, включенных в термин «антигенсвязывающий участок» антитела, включают в себя (i) фрагмент Fab, одновалентный фрагмент, состоящий из доменов VL, VH, CL и CH1; (ii) фрагмент F(ab')₂, двухвалентный фрагмент, содержащий два фрагмента Fab, связанных дисульфидным мостиком в шарнирной области; (iii) фрагмент Fd, состоящий из доменов VH и CH1; (iv) фрагмент Fv, состоящий из доменов VL и VH одного плеча антитела, (v) фрагмент dAt

(Ward et al. (1989) Nature 241:544-546), который состоит из домена VH, (vi) выделенный CDR и (vii) scFv, который состоит из двух доменов фрагмента Fv, VL и VH, соединенных синтетическим линкером с образованием одной белковой цепи, в которой области VL и VH соединяется с образованием одновалентных молекул. Другие формы одноцепочечных антител, таких как диатела, также охватываются термином «антитело» (см., например, Holliger et al. (1993) 90 PNAS U.S.A. 6444-6448; и Poljak et al. (1994) 2 Structure 1121-1123).

[0099] Более того, антитела и их антигенсвязывающие фрагменты могут быть получены с применением стандартных методов рекомбинантной ДНК, общеизвестных в данной области (см. Sambrook et al., 1989). Способы получения антител человека у трансгенных мышей также известны в данной области. Например, с применением технологии VELOCIMMUNE® (см., например, патент США 6,596,541, Regeneron Pharmaceuticals, VELOCIMMUNE®) или любого другого известного способа получения моноклональных антител, высокоаффинные химерные антитела к требуемому антигену изначально выделяют с вариабельной областью человека и константной областью мышей. Технология VELOCIMMUNE® включает создание трансгенной мыши, геном которой содержит вариабельные области тяжелой и легкой цепи человека, функционально связанные с эндогенными локусами константной области мыши, так что мышь продуцирует антитело, содержащее вариабельную область человека и константную область мыши, в ответ на антигенную стимуляцию. ДНК, кодирующую вариабельные области тяжелой и легкой цепей антитела, выделяют и функционально связывают с ДНК, кодирующей константные области тяжелой и легкой цепи человека. Затем ДНК экспрессируют в клетке, способной экспрессировать полностью антитело человека.

[00100] Предусмотрено, что термин «антитело человека» включает в себя антитела, имеющие вариабельные и константные области, происходящие из последовательностей иммуноглобулина зародышевой линии человека. mAb человека по изобретению могут включать в себя аминокислотные остатки, не кодируемые последовательностями иммуноглобулина зародышевой линии человека (например, мутации, введенные путем случайного или сайт-специфического мутагенеза *in vitro* или соматической мутации *in vivo*), например, в CDR и, в частности, CDR3. В то же время термин «человеческое антитело», используемый в данном документе, не предполагает включение mAb, в которых последовательности CDR, происходящие из зародышевой линии других видов млекопитающих (например, мыши), были привиты к FR последовательностям человека. Указанный термин включает в себя антитела, рекомбинантно продуцируемые в организме млекопитающего, не являющегося человеком, или в клетках млекопитающего, не являющегося человеком. Указанный термин не предназначен для включения антител, выделенных из или полученных в организме субъекта-человека.

[00101] Используемый в настоящем документе термин «примесь» может включать в себя любой нежелательный белок, присутствующий в биофармацевтическом продукте. Примесь может включать в себя связанные с процессом и продуктом примеси. Примесь может дополнительно представлять собой известную структуру, частично

охарактеризованную или неидентифицированную. Связанные с процессом примеси могут быть получены в результате производственного процесса и могут включать в себя три основные категории: полученные из клеточного субстрата, полученные из клеточной культуры и полученные в дальнейшем. Примеси, полученные из клеточного субстрата, включают в себя, помимо прочего, белки, полученные из организма-хозяина, и нуклеиновую кислоту (геномную, векторную или общую ДНК клетки-хозяина). Примеси, полученные из клеточной культуры, включают в себя, помимо прочего, индукторы, антибиотики, сыворотку и другие компоненты среды. Побочные примеси при последующей обработке, включают в себя, помимо прочего, ферменты, реагенты для химической и биохимической обработки (например, бромид цианогена, гуанидин, окислители и восстановители), неорганические соли (например, тяжелые металлы, мышьяк, неметаллический ион), растворители, носители, лиганды (например, моноклональные антитела) и другие выщелачиваемые вещества. Связанные с продуктом примеси (например, предшественники, определенные продукты разложения) могут представлять собой молекулярные варианты, возникающие в процессе производства и/или хранения, которые не обладают свойствами, сопоставимыми со свойствами требуемого продукта в отношении активности, эффективности и безопасности. Такие варианты могут потребовать значительных усилий при выделении и определении характеристик для идентификации типа модификации(-й). Связанные с продуктом примеси могут включать в себя усеченные формы, модифицированные формы и агрегаты. Усеченные формы образованы гидролитическими ферментами или химическими веществами, которые катализируют расщепление пептидных связей. Модифицированные формы включают в себя, помимо прочего, деамидированные, изомеризированные, несовпадающие S-S связанные, окисленные или измененные конъюгированные формы (например, гликозилирование, фосфорилирование). Модифицированные формы также могут включать в себя любую посттрансляционную форму модификации. Агрегаты включают в себя димеры и более высокие кратные требуемому продукту (Спецификации Q6B: Исследуемые процедуры и критерии приемлемости для BioTechnology/Biological Products, ICH август 1999 г., Департамент здравоохранения и социальных служб США).

[00102] Термин «низкомолекулярная (LMW) белковая примесь лекарственного средства» включает в себя, помимо прочего, предшественники, продукты разложения, усеченные виды, протеолитические фрагменты, включая фрагменты Fab, фрагменты Fc или тяжелой цепи, фрагменты лиганда или рецептора, виды H2L (2 тяжелые цепи и 1 легкая цепь), H2 (2 тяжелые цепи), HL (1 тяжелая цепь и 1 легкая цепь), HC (1 тяжелая цепь) и LC (1 легкая цепь). LMW белковая примесь лекарственного средства может представлять собой любой вариант, который представляет собой неполную версию белкового продукта, такого как один или более компонентов мультимерного белка. Белковая примесь лекарственного средства, примесь лекарственного средства или примесь продукта представляют собой термины, которые могут использоваться взаимозаменяемо по всему описанию. LMW примеси лекарственного средства или

продукта по существу считаются молекулярными вариантами со свойствами, такими как активность, эффективность и безопасность, которые могут отличаться от свойств требуемого продукта лекарственного средства.

[00103] Разложение белкового продукта является проблематичным при производстве белкового продукта лекарственного средства в системах клеточных культур. Например, протеолиз белкового продукта может происходить из-за высвобождения протеаз в среде клеточных культур. Средние добавки, такие как растворимые источники железа, добавляемые для ингибирования металлопротеаз, или ингибиторы протеаз серина и цистеина были реализованы в клеточной культуре для предотвращения разложения (Clincke, M.-F., et al, BMC Proc. 2011, 5, P115). С-концевые фрагменты могут быть расщеплены во время производства из-за карбоксилпептидаз в клеточной культуре (Dick, LW et al, Biotechnol Bioeng 2008; 100:1132-43).

[00104] Термин «высокомолекулярная (HMW) белковая примесь лекарственного средства» включает в себя, помимо прочего, тримеры mAb и димеры mAb. Виды HMW можно разделить на две группы: 1) мономер с дополнительными легкими цепями (виды H2L3 и H2L4) и 2) мономер плюс комплексы фрагментов Fab. Кроме того, после лечения ферментативным расщеплением IdeS образуются различные димеризованные фрагменты (Fab2-Fab2, Fc-Fc и Fab2-Fc).

[00105] Используемый в настоящем документе термин «гликопептид/гликопротеин» представляет собой модифицированный пептид/белок, во время или после их синтеза, с ковалентно связанными углеводами или гликаном. В определенных вариантах осуществления гликопептид получают из моноклонального антитела, например, из протеазного гидролизата моноклонального антитела.

[00106] Используемый в настоящем документе термин «гликан» представляет собой соединение, содержащее одно или более из сахарных единиц, которые обычно включают в себя глюкозу (Glc), галактозу (Gal), маннозу (Man), фукозу (Fuc), N-ацетилгалактозамин (GalNAc), N-ацетилглюкозамин (GlcNAc) и N-ацетилнейраминую кислоту (NeuNAc) (Frank Kjeldsen, et al. Anal. Chem. 2003, 75, 2355-2361). Функциональная группа гликана в гликопротеине, таком как моноклональное антитело, является важным символом для идентификации его функции или местоположения клеток. Например, специфическое моноклональное антитело модифицировано специфической функциональной группой гликана.

[00107] Используемый в настоящем документе термин «образец» включает в себя по меньшей мере молекулу аналита, например гликопептид, например, полученный из моноклонального антитела, который подвергают манипуляции в соответствии со способами изобретения, включая, например, разделение, анализ, экстрагирование, концентрацию или профилирование.

[00108] Используемые в настоящем документе термины «анализ» или «анализирование» используются взаимозаменяемо и относятся к любому из различных способов разделения, обнаружения, выделения, очистки, солиubilизации, обнаружения

и/или определения характеристик интересующих молекул. Примеры включают в себя, помимо прочего, хроматографию, твердофазное экстрагирование, твердофазное микроэкстрагирование, электрофорез, масс-спектрометрию, например жидкостную хроматографию, например с высокой эффективностью, например, обращенно-фазовую, нормально-фазовую или размерно-эксклюзионную, ионно-парную жидкостную хроматографию, жидкостно-жидкостную экстракцию, например ускоренное жидкостное экстрагирование, сверхкритическое жидкостное экстрагирование, микроволновое экстрагирование, мембранное экстрагирование, экстрагирование по Сокслету, осаждение, осветление, электрохимическое обнаружение, окрашивание, элементный анализ, расщепление по Эдмунду, ядерно-магнитный резонанс, инфракрасный анализ, проточно-инжекционный анализ, капиллярную электрохроматографию, ультрафиолетовое обнаружение и их комбинации.

[00109] Используемый в настоящем документе термин «профилирование» относится к любому из различных способов анализа, которые используются в комбинации для обеспечения содержания, состава или характеристического соотношения соединений, таких как белки.

[00110] Используемый в настоящем документе термин «расщепление» относится к гидролизу одной или более пептидных связей белка. Существует несколько подходов для осуществления расщепления белка в образце с применением соответствующего гидролизующего агента, например ферментативного расщепления или неферментативного расщепления. Используемый в настоящем документе термин «гидролизующий агент» относится к любой одной или комбинации большого количества различных агентов, которые могут выполнять расщепление белка. Не имеющие ограничительного характера примеры гидролизующих агентов, которые могут осуществлять ферментативное расщепление, включают в себя трипсин, эндопротеиназу Arg-C, эндопротеиназу Asp-N, эндопротеиназу Glu-C, внешнюю мембранную протеазу T (OmpT), иммуноглобулин-разлагающий фермент *Streptococcus pyogenes* (IdeS), химотрипсин, пепсин, термоллизин, папаин, проназу и протеазу из *Aspergillus Saitoi*. Не имеющие ограничительного характера примеры гидролизующих агентов, которые могут осуществлять неферментативное расщепление, включают в себя применение высокой температуры, микроволнового излучения, ультразвука, высокого давления, инфракрасного излучения, растворителей (не имеющие ограничительного характера примеры представляют собой этанол и ацетонитрил), расщепления иммобилизованного фермента (IMER), иммобилизованных на магнитных частицах ферментов и иммобилизованных на кристалле ферментов. В отношении последнего обзора с рассмотрением доступных методов расщепления белка см. Switazar et al., «Protein Digestion: An Overview of the Available Techniques and Recent Developments» (*J. Proteome Research* 2013, 12, 1067-1077). Один или комбинация гидролизующих агентов могут расщеплять пептидные связи в белке или полипептиде специфичным для последовательности образом, создавая прогнозируемый набор более коротких пептидов.

[00111] Доступны несколько подходов, которые можно применять для расщепления белка. Один из широко принятых способов расщепления белков в образце включает в себя применение протеаз. Доступны много протеаз, и каждая из них имеет свои собственные характеристики с точки зрения специфичности, эффективности и оптимальных условий расщепления. Протеазы относятся как к эндопептидазам, так и к экзопептидазам, которые классифицируются на основе способности протеазы расщеплять неконцевые или концевые аминокислоты внутри пептида. Альтернативно, протеазы также относятся к шести различным классам - аспарагиновым, глутаминовым и металлопротеазам, цистеиновым, сериновым и треониновым протеазам, как классифицировано механизмом катализа. Термины «протеаза» и «пептидаза» используются взаимозаменяемо для обозначения ферментов, которые гидролизуют пептидные связи. Протеазы также могут быть классифицированы как специфические и неспецифические протеазы. Используемый в настоящем документе термин «специфическая протеаза» относится к протеазе, обладающей способностью расщеплять пептидный субстрат на специфической боковой аминокислотной цепи пептида. Используемый в настоящем документе термин «неспецифическая протеаза» относится к протеазе, обладающей пониженной способностью расщеплять пептидный субстрат на специфической боковой аминокислотной цепи пептида. Предпочтение расщепления может быть определено на основе соотношения количества конкретной аминокислоты в качестве сайта расщепления к общему количеству расщепленных аминокислот в белковых последовательностях.

[00112] Белок необязательно может быть получен до определения характеристик. В некоторых примерах осуществления получение белка включает в себя стадию расщепления белка. В некоторых конкретных примерах осуществления получение белка включает в себя стадию расщепления белка, причем расщепление белка может быть выполнено с применением трипсина.

[00113] В некоторых примерах осуществления получение белка может включать в себя стадию денатурирования белка, восстановления белка, забуферивания белка и/или обессоливания образца перед стадией расщепления белка. Эти стадии могут быть выполнены при необходимости любым подходящим способом.

[00114] Используемый в настоящем документе термин «хроматография» относится к методу обработки для разделения компонентов, или растворенных веществ, смеси на основе относительных количеств каждого растворенного вещества, распределенного между движущимся потоком текучей среды, называемым подвижной фазой, и смежной неподвижной фазой. Подвижная фаза может представлять собой либо жидкость, либо газ, тогда как неподвижная фаза представляет собой твердое вещество или жидкость.

[00115] Используемый в настоящем документе термин «жидкостная хроматография» относится к процессу, в котором химическая смесь, переносимая жидкостью, может быть разделена на компоненты в результате дифференциального распределения химических объектов при их протекании вокруг или над неподвижной

жидкостью или твердой фазой. Не имеющие ограничительного характера примеры жидкостной хроматографии включают в себя обращенно-фазовую жидкостную хроматографию, ионообменную хроматографию, размерно-экслюзионную хроматографию, аффинную хроматографию и гидрофобную хроматографию.

[00116] Используемый в настоящем документе термин «многомерный инструмент» относится к статистическому инструменту, который применяет множество переменных для прогнозирования результатов. Многомерный инструмент может позволять исследовать, анализировать и/или интерпретировать данные. Инструмент может облегчать ввод данных путем раскрытия тенденций и кластеров, анализа вариаций процесса, идентификации параметров и/или прогнозирования конечного качества продукта. В некоторых примерах многомерный инструмент является коммерчески доступным, например SIMCA (umetrics, Умеа, Швеция).

[00117] Используемый в настоящем документе термин «покрытие белковой последовательности» относится к процентному содержанию белковой последовательности, покрытой идентифицированными пептидами. Процентное содержание покрытия может быть рассчитано путем деления количества аминокислот во всех обнаруженных пептидах на общее количество аминокислот во всей белковой последовательности.

[00118] Используемый в настоящем документе термин «база данных» относится к биоинформационным инструментам, которые обеспечивают возможность поиска неинтерпретированных спектров MS-MS по всем возможным последовательностям в базе(-ах) данных. Не имеющие ограничительного характера примеры таких инструментов представляют собой Mascot (www.matrixscience.com), Spectrum Mill (www.chem.agilent.com), PLGS (www.waters.com), PEAKS (www.bioinformaticssolutions.com), Proteinpilot (download.appliedbiosystems.com/proteinpilot), Phenyx (<http://www.phenyx-ms.com>), Sorcerer (www.sagenresearch.com), OMSSA (www.pubchem.ncbi.nlm.nih.gov/omssa/), X!Tandem (www.thegpm.org/TANDEM/), Protein Prospector (<http://prospector.ucsf.edu/prospector/mshome.htm>), Byonic (www.proteinmetrics.com/products/byonic) или Sequest (fields.scripps.edu/sequest).

Общее описание

[00119] Из вышеизложенного следует понимать, что существует потребность в улучшенных способах и системах для улучшения очистки белков, включая очистку антител. Описанное изобретение удовлетворяет данную потребность. В настоящем документе описаны способы с применением моделирования OPLS, включая примеры приемлемых (без ухудшения состояния) и неприемлемых (с ухудшением состояния) УФ-хроматограмм, таких как УФ-, инфракрасные (ИК)- или Ramen-кривые, для обеспечения раннего обнаружения разрушения колонки и количественного анализа УФ-сигнала в хроматограмме. В некоторых вариантах осуществления в способах применяют УФ-моделирование OPLS. Описанные способы объединяют технологические знания,

полученные специалистами в предметной области (SME) с уравнениями и процедурами для создания хроматограмм для создания комплексных наборов данных в диапазоне от надежных до приемлемых, до постепенных отказов, до катастрофических (см. Фиг. 1А). В описанных способах применяют OPLS для решения многих проблем, таких как автоматизация и оценка комплексных данных, что ранее не было возможно. Описанные способы обеспечивают контроль процесса посредством применения многомерного инструмента (например, SIMCA) для прогнозирования разрушения колонки. Описанные модели OPLS могут быть построены без статистических разрушений колонок и могут быть автоматизированы (например, с применением сценариев Python), исключая ручной процесс для выполнения наложений и анализа. Отказы процедурно сгенерированных операций с применением математических уравнений и протоколов преодолевают слабые стороны ограниченных наборов статистических данных, не содержащих отказы. Описанные способы могут обеспечивать количественное измерение хроматограммы и характеристик колонки и обеспечивать динамическую идентификацию вариации. Таким образом, описанные способы обеспечивают согласованность и надежность процесса, а также постоянные характеристики оборудования, устраняя при этом недостатки, связанные с существующими способами.

[00120] В вариантах осуществления в описанных способах применяют OPLS. Регрессия OPLS анализирует взаимосвязи между двумя блоками данных путем расширений регрессии анализа основных компонентов (PCA; Principle Component Analysis). Например, создается пространство данных для матриц X и Y. Первый компонент соответствует линии в пространстве X и Y таким образом, что корреляция между проекциями X и Y максимально увеличена. Второй компонент является ортогональным первому компоненту в пространстве X. Первый компонент является прогнозирующим и максимально увеличивает ковариацию между X и Y, тогда как второй компонент и далее являются ортогональными и представляют структурированную вариацию X, которая не связана с Y (Фиг. 1В). OPLS-DA представляет собой модель, в которой качественные категориальные переменные (например, хорошая/удовлетворительная или плохая/неудовлетворительная) являются назначенными количественными значениями, так что может быть построена модель OPLS. На Фиг. 1С показано схематическое изображение, иллюстрирующее прогноз OPLS.

[00121] В вариантах осуществления способ контроля характеристик колоночной хроматографии включает в себя получение одной или более хроматограмм в ультрафиолетовом спектре (УФ-хроматограмм), сгенерированных системой хроматографии во время очистки и/или разделения образца; и анализ указанных одной или более УФ-хроматограмм с помощью модели OPLS, таким образом обеспечивая обнаружение ухудшения состояния колонки до разрушения колонки и количественный анализ УФ-сигнала в указанных одной или более УФ-хроматограмм.

[00122] В вариантах осуществления способ контроля характеристик колоночной хроматографии включает в себя создание УФ-модели OPLS. На Фиг. 1D предоставлена

блок-схема, иллюстрирующая иллюстративный протокол создания УФ-модели OPLS. Как проиллюстрировано на Фиг. 1D, создание УФ-модели OPLS включает в себя стадию 102, идентификацию процесса или однократной операции создаваемой УФ-модели OPLS. В некоторых вариантах осуществления это конкретная хроматографическая однократная операция для определенного продукта/молекулы, например аффинность белка, ионный обмен, стадия гидрофобного взаимодействия или размерно-эксклюзионной хроматографии либо для связывания и элюирования, либо для проточной конструкции для конкретной молекулы. В вариантах осуществления по меньшей мере часть УФ-сигнала в хроматограмме подается в модель. Некоторые части хроматограммы не полезны для определения целостности колонки. Например, применяют по меньшей мере один блок или секцию хроматограммы. В некоторых вариантах осуществления, если метод колонки однократной операции представляет собой связывание и элюирование, то блок элюирования применяют для анализа (например, когда связанный продукт элюируется из колонки). На Фиг. 2 представлен пример хроматограммы связывания с элюированием с выделенным блоком элюирования. В некоторых вариантах осуществления, если метод колонки однократной операции является проточным, будет проанализирована точка, в которой сбор начинается во время отрыва УФ-сигнала. На Фиг. 3 представлен пример проточной хроматограммы с выделенным УФ-отрывом.

[00123] Способ, проиллюстрированный на Фиг. 1D для создания УФ-модели OPLS, дополнительно включает в себя стадию 200, сбор исходных данных для УФ-кривой. В некоторых вариантах осуществления сбор исходных данных для УФ-кривой включает в себя сбор исходных данных множества циклов, партий и/или прогонов для выбранной однократной операции, например значений УФ-поглощения в соответствующих объемах. Например, исходные данные для УФ и регистрационного журнала из программного обеспечения хроматографии (например, Unicorn) экспортируют, например, в электронный файл хранения (например, файл Microsoft Excel). Регистрационный журнал представляет собой трекер, который идентифицирует, какие события и в каком объеме произошли во время хроматограммы, и идентифицирует, куда следует извлечь блок данных для ввода в модель (например, блок элюирования). Затем процесс можно повторять для каждого прогона/цикла хроматографии. В вариантах осуществления прогон/цикл исходных данных хранится, например, в отдельном файле (например, отдельных файлах Microsoft Excel), как показано на Фиг. 4.

[00124] В вариантах осуществления иллюстративный способ включает в себя стадию 106, нормализацию и совмещение исходных данных, например нормализацию УФ-значений и совмещение объемов. Например, данные нормализуют для удаления любой вариации различий величин (от функциональной возможности УФ-метра) в исходном сигнале. В вариантах осуществления исходные данные УФ-излучения и регистрационного журнала импортируют, данные нормализуют и форматируют для импорта в многомерный инструмент (например, SIMCA), например, с применением

макроса. В вариантах осуществления стадии, выполняемые макросом, могут включать в себя следующее:

1. Сначала макрос предлагает выбрать файлы, по которым будет работать макрос (например, какие прогоны будут использованы для построения модели);
2. Далее макрос находит участок хроматограммы, который будет анализироваться моделью (например, блок «Элюирование»);
3. Макрос находит максимальные и минимальные значения УФ излучения и нормализует значения УФ-излучения от -1 до 1;
4. Макрос нормализует объем с точки зрения объемов колонки (CV);
5. Макрос извлекает нормализованные УФ-данные с указанным интервалом CV, чтобы все УФ-данные были совмещены (например, извлекает значение УФ-сигнала каждые 0,02 CV в блоке «Элюирование»);
6. Макрос вставляет нормализованные данные в таблицу, отформатированную для импорта Simca; и
7. Макрос повторяет стадии 2-6, если выбрано более одного прогона.

В вариантах осуществления объема колонки (CV) оптимизированы в виде интервала объема для максимального увеличения разрешения данных и сведения к минимуму требуемых вычислительных ресурсов. В одном варианте осуществления используют 0,02 CV. На Фиг. 5-1 и Фиг. 5-2 приведен пример нормализации данных в соответствии с вариантами осуществления, описанными в настоящем документе.

[00125] В вариантах осуществления способ включает в себя оценку события разрушения колонки в данных перед генерированием кривых. Например, если имеются эмпирические примеры неудачных хроматограмм, то могут не потребоваться искусственно сгенерированные хроматограммы. Если имеется достаточное количество примеров (например, 7 или более, например 8, 9, 10, 11, 12, 13, 14, 15 или более) неудачных хроматограмм, то эти хроматограммы служат для неприемлемого набора данных, и искусственное генерирование кривой не требуется.

[00126] В вариантах осуществления способ необязательно включает в себя стадию 108, создание искусственно сгенерированных кривых на основании нормализованных данных для получения примеров модели OPLS нежелательных хроматограмм. Например, стадию 108 выполняют, если отсутствуют примеры неприемлемых хроматограмм. В вариантах осуществления, если хроматограмма предназначена для колонки связывания и элюирования, применяют инструмент для генерирования пиков элюирования, такой как инструмент Microsoft Excel для генерирования пиков элюирования. В вариантах осуществления, если хроматограмма предназначена для проточной колонки, применяют инструмент Excel для генерирования проточных кривых.

[00127] В некоторых вариантах осуществления создание искусственно сгенерированных кривых включает в себя определение того, какие прогоны применять для генерирования кривых. Например, по существу около 10-15 прогонов представляют собой достаточное количество хроматограмм для оценки. Если набор данных является

большим (например, более 100 прогонов), то выбирают прогоны таким образом, чтобы зафиксировать изменчивость процесса (например, выбирают только приемлемые хроматограммы). В различных вариантах осуществления создание искусственно сгенерированных кривых дополнительно включает в себя копирование данных на вкладку генерирования в F-Q и копирование данных для одного из прогонов в колонку E (например, на диаграмме хроматограмма отображается первым цветом, например синим, на фигуре рядом с данными, тогда как математически сгенерированная хроматограмма отображается вторым цветом, например красным). В некоторых вариантах осуществления способ создания искусственно сгенерированных кривых включает в себя совмещение кривой, сгенерированной на основании модели, для данных реального прогона. Например, совмещение может продолжаться до тех пор, пока не будет прекращено уменьшение в клетке X1 значение среднеквадратичной ошибки (RMSE), что указывает на то, что модель была подогнана к хроматограмме. Данные параметров уравнения копируют для кривой, и такие значения будут применять для процедурного генерирования новых хроматограмм. Одно или более из предшествующих действий могут быть повторены для выбранных данных прогона. Например, прогоны выбирают с вариабельностью, поскольку не требуется подгонять множество прогонов, если они являются относительно зеркальными отражениями.

[00128] В различных вариантах осуществления создание искусственно сгенерированных кривых для нормализованных данных включает в себя генерирование хроматограмм, которые являются типичными для ухудшения состояния. Например, это может быть выполнено с применением компьютерной программы, такой как Microsoft Excel, которая копирует средние значения и значения среднеквадратичного отклонения для каждого параметра уравнения в таблицу в столбцах X и Y. В вариантах осуществления для генерирования пиков элюирования используются четыре термина в математической модели: (1) $Tm1$: определяет, где происходит пиковый отрыв (например, уменьшение $tm1$ может сдвинуть начало пика влево, а его увеличение может сдвинуть его вправо); (2) $S1$: определяет, насколько крутым является пиковый отрыв (например, уменьшение $s1$ приводит к тому, что пиковый отрыв становится более резким, в то время как увеличение делает его более широким); (3) $Tm2$: определяет, где находится конец пика (например, уменьшение $tm2$ сдвигает конец пика влево, а увеличение $tm2$ сдвигает конец пика вправо); и $S2$: определяет, насколько крутым является падение пика (например, уменьшение $s2$ делает конец пика более резким, в то время как увеличение $s2$ делает конец пика более широким). В вариантах осуществления могут применять дополнительные условия для подгонки максимума пика, например, если верхняя часть пика элюирования не является плоской или если применяемый УФ-датчик требует дополнительных входных данных. В вариантах осуществления параметры уравнения изначально установлены равными средним значениям параметров из вкладки генерирования, а значения среднеквадратичного отклонения (SD) изначально установлены равными 0. Изменение этих значений изменит соответствующий параметр уравнения на

количество среднеквадратичных отклонений, введенных в клетку (например, ввод 3 в клетку SD рядом с s1 увеличит значение s1 на 3 SD). В вариантах осуществления выполняется доступ к каждому параметру, и значение SD либо увеличивается, либо уменьшается для создания пиков элюирования, свидетельствующих об ухудшении состояния, например расширении (увеличении tm_2), затягивании (увеличении s_2); фронтальные и/или биомодальные пики представляют собой другие распространенные показатели ухудшения состояния, которые также можно моделировать. Выполняют действие, позволяющее создавать примеры незначительных, умеренных и/или значительных вариаций. В некоторых вариантах осуществления с применением вышеупомянутого процесса создается 10-15 нежелательных хроматограмм, например создается одно или более незначительных отклонений, одно или более умеренных и одно или более примеров значительных вариаций. В вариантах осуществления нежелательные хроматограммы включают в себя большее количество примеров незначительных отклонений по сравнению с примерами умеренных и значительных отклонений. В вариантах осуществления нежелательные хроматограммы включают в себя большее количество примеров умеренных отклонений по сравнению с примерами значительных отклонений. Например, в одном варианте осуществления созданные нежелательные хроматограммы включают в себя 7 незначительных отклонений, 5 умеренных отклонений и 3 в качестве примеров значительных отклонений. На Фиг. 6-1, 6-2, 6-3, 7, 8, 9-1, 9-2, 10-1 и 10-2 проиллюстрировано создание искусственно сгенерированных кривых на основе нормализованных данных в соответствии с вариантами осуществления, описанными в настоящем документе.

[00129] В вариантах осуществления способ дополнительно включает в себя стадию 110, классификацию данных и форматирование данных для импорта в многомерный инструмент, такой как SIMCA, для определения того, какие данные являются приемлемыми и неприемлемыми для обучения модели. На Фиг. 11-1 и Фиг. 11-2 показана иллюстративная классификация и форматирование данных для импортирования SIMCA в соответствии с вариантами осуществления, описанными в настоящем документе. Например, в некоторых вариантах осуществления классификация и форматирование данных для импортирования в многомерный инструмент включает в себя (1) осуществление доступа к файлу, применяемому для нормализации данных; (2) копирование данных для искусственно сгенерированных хороших и плохих кривых после уже импортированных данных; (3) классификацию каждого прогона/цикла с разделением на две группы: приемлемую или «хорошую» группу, которая содержит хроматограммы, не показывающие признаков ухудшения состояния колонки, и неприемлемую или «плохую» группу, которая содержит хроматограммы, имеющие показатели ухудшения состояния колонки, такие как затягивание, расширение пика или другие показатели, известные специалистам в данной области, указывающие на ухудшение состояния колонки; и (4) обеспечение дополнительных категорий для классификации прогонов, таких как номер заполнения колонки или номер цикла смолы, что облегчает возможность

организации и считывания данных с помощью многомерного инструмента, такого как SIMCA (например, сгенерированные группы могут быть окрашены и отсортированы в SIMCA).

[00130] В вариантах осуществления способ включает в себя импорт данных в многомерную платформу, такую как SIMCA (стадия 112), и генерирование модели OPLS (стадия 114). На Фиг. 12А, 12В, 13А, 13В-1, 13В-2, 14А, 14В, 15А, 15В-1, 15В-2 и 16 проиллюстрирован протокол для импорта данных в SIMCA в соответствии с вариантами осуществления, описанными в настоящем документе. На Фиг. 17, 18, 19 и 20 проиллюстрирован протокол генерирования модели OPLS в соответствии с вариантами осуществления, описанными в настоящем документе.

[00131] После генерирования модели OPLS способ включает в себя оптимизацию модели OPLS (стадия 116), подтверждение и тестирование (стадия 118) и применение модели (стадия 120). В некоторых вариантах осуществления применение модели может включать в себя классификацию новых данных как приемлемых или неприемлемых. В вариантах осуществления полученную в результате оптимизированную модель OPLS далее дополняют, например, путем увеличения, уменьшения и/или изменения количества и/или содержания математически сгенерированных входных данных хроматограммы. Например, если модель OPLS является неудовлетворительной после оптимизации настроек OPLS, ее далее дополняют путем увеличения, уменьшения и/или изменения количества и/или содержания математически сгенерированных входных данных хроматограммы. В вариантах осуществления оптимизация модели OPLS является итеративной с математически сгенерированными примерами отказа хроматограмм, описанными в настоящем документе. На Фиг. 21, 22А, 22В, 23А, 23В, 24А, 24В, 25-1 и 25-2 показана иллюстративная оптимизация модели OPLS в соответствии с вариантами осуществления, описанными в настоящем документе. В вариантах осуществления оптимизация OPLS включает в себя определение того, существует ли другой набор параметров создания модели, который соответствует лучшей модели для данных. Значение Q^2 представляет прогнозирующую способность модели (например, чем выше Q^2 , тем выше прогнозирующая способность модели). В различных вариантах осуществления значение Q^2 , составляющее 0,7 или более, считается приемлемым. В некоторых вариантах осуществления дополнительные модели создают путем сравнения значений Q^2 для определения того, какая модель более способна прогнозировать новые данные. Кроме того, с помощью многомерного инструмента, такого как SIMCA, можно выполнять дополнительные преобразования данных для анализа производной или преобразованной версии импортированных данных. В вариантах осуществления один или более из следующих параметров корректируют для создания дополнительных моделей OPLS: (1) Производный уровень (например, 0 или непроизводные данные, 1 или первая производная, 2 или вторая производная и т. д.); (2) Уровень сглаживания (например, при взятии производных или для дополнительного сглаживания данных); и/или (3)

Исключение наблюдений (например, неисправность УФ-метра на протяжении хроматограммы).

[00132] В вариантах осуществления оптимизированную УФ-модель OPLS анализируют для определения прогнозирующей способности модели. Например, многомерный инструмент, такой как SIMCA, применяют для тестирования прогнозирующей способности оптимизированной УФ-модели OPLS. В вариантах осуществления определяют перестановки, которые указывают на статистическую значимость значений R^2 и Q^2 , путем нахождения эталонных распределений значений R^2 и Q^2 в результате тестирования перестановок переменной Y . Если модель представляет собой удовлетворительную модель, эталонные значения R^2 и Q^2 будут иметь более низкие значения, чем значения модели. В вариантах осуществления генерируют график показателей CV, который показывает компонент перекрестного подтверждения для графика обычных показателей. График показателей CV, который близко соответствует графику обычных показателей, указывает на то, что модель является удовлетворительной.

[00133] На Фиг. 26А, 26В, 27А, 27В, 28, 29-1, 29-2, 30, 31-1 и 31-2 показано иллюстративное применение созданной модели OPLS для классификации новых хроматограмм как приемлемых или неприемлемых в соответствии с вариантами осуществления, описанными в настоящем документе.

[00134] В вариантах осуществления способ дополнительно включает в себя предоставление образца в систему хроматографии перед получением одной или более УФ-хроматограмм, генерируемых системой хроматографии во время очистки и/или разделения образца.

[00135] Хотя в описании в настоящем документе подробно описано применение описанной модели OPLS с УФ-спектрофотометрией, предполагается, что описанная модель и способы могут быть применены с любой формой спектрофотометрии, которая контролирует выход колонки для элюата, включая, помимо прочего, для RAMAN или ИК.

[00136] В некоторых вариантах осуществления система хроматографии, способная разделять образец для очистки и/или разделения компонентов образца, включает в себя систему жидкостной хроматографии. В некоторых вариантах осуществления система представляет собой систему хроматографии, которая представляет собой систему гидрофобной хроматографии, систему обращенно-фазовой жидкостной хроматографии, систему ионообменной хроматографии, систему размерно-эксклюзионной хроматографии, систему аффинной хроматографии или систему хроматографии гидрофильного взаимодействия.

[00137] В некоторых вариантах осуществления температуру хроматографической колонки можно поддерживать на постоянном уровне температуры в течение всего прогона хроматографии, например, с применением промышленного нагревателя колонки. В некоторых вариантах осуществления колонка поддерживается при температуре от около 18 °C до около 70 °C, например, от около 30 °C до около 60 °C, от около 40 °C до около 50 °C, например при около 20 °C, около 25 °C, около 30 °C, около 35 °C, около 40 °C, около

45 °С, около 50 °С, около 55 °С, около 60 °С, около 65 °С или около 70 °С. В некоторых вариантах осуществления температура колонки составляет около 40 °С. В некоторых вариантах осуществления время прогона может составлять от около 15 мин до около 240 минут, например от около 20 мин до около 70 мин, от около 30 мин до около 60 мин, от около 40 мин до около 90 мин, от около 50 мин до около 100 мин, от около 60 мин до около 120 мин, от около 50 мин до около 80 мин.

[00138] В некоторых вариантах осуществления подвижная фаза представляет собой водную подвижную фазу. Типичная водная подвижная фаза содержит 208 мМ ацетата натрия и 10 мМ бикарбоната аммония. УФ-кривые обычно регистрируют при 215 нм и 280 нм.

[00139] В некоторых иллюстративных вариантах осуществления подвижная фаза, применяемая для элюирования белка, может представлять собой подвижную фазу, которая может быть совместима с масс-спектрометром.

[00140] В некоторых примерах осуществления подвижная фаза, применяемая в устройстве для жидкостной хроматографии, может содержать воду, ацетонитрил, трифторуксусную кислоту, муравьиную кислоту или их комбинацию.

[00141] В некоторых примерах осуществления подвижная фаза для производственных операций может иметь скорость потока, изменяющуюся в течение операции и в зависимости от операции от 60 л/ч до 1800 л/ч.

[00142] В некоторых вариантах осуществления образец представляет собой белок или среду клеточных культур, содержащую белок, иллюстративные белки, включая, помимо прочего, антитело, слитый белок, рекомбинантный белок или их комбинацию.

[00143] В некоторых вариантах осуществления антитело представляет собой биспецифическое антитело, фрагмент антитела или мультиспецифическое антитело.

[00144] В некоторых примерах осуществления антитело представляет собой моноклональное антитело, такое как, помимо прочего, моноклональное антитело изотипа IgG1, IgG2, IgG3, IgG4 или смешанного изотипа.

[00145] В некоторых примерах осуществления белок представляет собой терапевтический белок.

[00146] В некоторых примерах осуществления белок может представлять собой белок иммуноглобулина.

[00147] В одном примере осуществления белок может представлять собой вариант белка.

[00148] В одном примере осуществления белок может представлять собой посттрансляционно модифицированный белок.

[00149] В одном примере осуществления посттрансляционно модифицированный белок может быть образован путем расщепления, удлинений N-конца, разложения белка, ацилирования N-конца, биотинилирования, амидирования C-конца, окисления, гликозилирования, йодирования, ковалентного присоединения простетических групп, ацетилирования, алкилирования, метилирования, аденилирования, АДФ-

рибозилирования, ковалентных поперечных связей внутри полипептидных цепей или между ними, сульфирования, пренилирования, модификаций, зависящих от витамина С, модификаций, зависящих от витамина К, глутамилирования, глицилирования, гликозилирования, дегликозилирования, изопренилирования, липоилирования, фосфопантетеинилирования, фосфорилирования, сульфатирования, цитруллинирования, дезамидирования, образования дисульфидных мостиков, протеолитического расщепления, ISG-илирования, SUMO-илирования или убиквитинирования (ковалентного связывания с белком убиквитином).

[00150] В одном примере осуществления посттрансляционно модифицированный белок может быть образован при окислении белка.

[00151] В вариантах осуществления описанные способы применяют для контроля ухудшения состояния колонки из-за изменения состояния заполнения колонки, накопления загрязняющих компонентов, прохождения через колонку, блокировки микрочастиц, десорбции из твердой фазы или их комбинации. В вариантах осуществления описанные способы обнаруживают ухудшение состояния колонки перед разрушением колонки. В вариантах осуществления описанные способы обнаруживают неизбежное ухудшение состояния колонки перед признаком ухудшения состояния колонки, таким как повышенное давление в колонке, уменьшенные теоретические пластины, сокращенное время удерживания, плохую форму пика и/или снижение разрешения.

[00152] Предполагается, что способы, описанные в настоящем документе, могут быть выполнены с помощью программного обеспечения, аппаратного обеспечения или обоих вычислительной среды, например одного или более вычислительных устройств. Например, вычислительные устройства включают в себя серверные компьютеры, настольные компьютеры, портативные компьютеры, ноутбуки, переносные устройства, нетбуки, планшетные устройства, мобильные устройства и другие типы вычислительных устройств.

[00153] На Фиг. 41 представлена иллюстративная вычислительная среда 200 для осуществления различных аспектов способов, описанных в настоящем документе, включая создание и/или применение OPLS для контроля способов контроля характеристик хроматографии. Вычислительная среда 200 не предназначена для указания какого-либо ограничения объема применения или функциональных возможностей, поскольку технологии могут быть реализованы в различных вычислительных средах общего назначения или специального назначения. Например, описанная технология может быть реализована с применением вычислительного устройства, содержащего блок обработки, запоминающее устройство и устройство хранения, сохраняющего исполняемые компьютером команды, реализующие способы, описанные в настоящем документе. Описанная технология также может быть реализована с другими конфигурациями компьютерных систем, включая портативные устройства, многопроцессорные системы, микропроцессорную или программируемую бытовую электронику, сетевые ПК, мини-компьютеры, мейнфреймы-компьютеры, набор систем

клиент/сервер и т. п. Описанная технология также может быть реализована на практике в распределенных вычислительных средах, в которых задачи выполняются удаленными устройствами обработки, связанными через сеть связи. В распределенной вычислительной среде программные модули могут располагаться как в локальных, так и в удаленных запоминающих устройствах.

[00154] Со ссылкой на Фиг. 41, вычислительная среда 200 содержит по меньшей мере один блок 210 обработки, соединенный с запоминающим устройством 220. На Фиг. 41 эта базовая конфигурация 230 обведена пунктирной линией. Блок 210 обработки выполняет исполняемые компьютером команды и может быть реальным или виртуальным процессором. В многопроцессорной системе множество блоков обработки выполняют исполняемые компьютером команды для увеличения вычислительной мощности. Запоминающее устройство 220 может представлять собой энергозависимое запоминающее устройство (например, регистры, кэш-память, ОЗУ), энергонезависимое запоминающее устройство (например, ПЗУ, ЭСППЗУ, флэш-память и т. д.) или определенную комбинацию обоих. Запоминающее устройство 220 может хранить программное обеспечение 280, реализующее любую из описанных в настоящем документе технологий.

[00155] Вычислительная среда может иметь дополнительные функции. Например, вычислительная среда 200 включает в себя устройство 240 хранения данных, одно или более устройств 250 ввода, одно или более устройств 260 вывода и одно или более соединений 270 связи. Механизм взаимных соединений (не показан), такой как шина, контроллер или сеть, взаимно соединяет компоненты вычислительной среды 200. Как правило, программное обеспечение операционной системы (не показано) обеспечивает операционную среду для другого программного обеспечения, выполняемого в вычислительной среде 200, и координирует действия компонентов вычислительной среды 200.

[00156] Устройство 240 хранения данных может быть съемным или несъемным и включает в себя магнитные диски, магнитные ленты или кассеты, CD-ROM, CD-RW, DVD или любые другие машиночитаемые носители, которые можно применять для хранения информации и к которым можно получить доступ внутри вычислительной среды 200. В устройстве 240 хранения данных может храниться программное обеспечение 280, содержащее команды для любой из описанных в настоящем документе технологий.

[00157] Устройство(-а) 250 ввода может(-гут) представлять собой сенсорное устройство ввода, такое как клавиатура, мышь, перо или шаровой манипулятор, устройство голосового ввода, сканирующее устройство или другое устройство, обеспечивающее ввод в вычислительную среду 200. Для аудио устройство(-а) 250 ввода может(-гут) представлять собой звуковую карту или аналогичное устройство, которое принимает аудиоввод в аналоговой или цифровой форме, или считыватель CD-ROM, который предоставляет аудиосэмплы в вычислительную среду. Устройство(-а) 260 вывода

может(-гут) представлять собой дисплей, принтер, динамик, устройство записи компакт-дисков или другое устройство, обеспечивающее вывод из вычислительной среды 200.

[00158] Соединение(-я) 270 связи обеспечивает(-ют) связь через механизм связи с другим вычислительным объектом. Механизм связи передает информацию, такую как исполняемые компьютером команды, аудио/видео или другую информацию или другие данные. В качестве примера, а не ограничения, механизмы связи включают в себя проводные или беспроводные методы, реализованные с помощью электрического, оптического, радиочастотного (РЧ), инфракрасного, акустического или другого носителя.

[00159] Методы, описанные в настоящем документе, могут быть описаны в общем контексте исполняемых компьютером команд, таких как команды, включенные в программные модули, которые выполняются в вычислительной среде на целевом реальном или виртуальном процессоре. По существу программные модули включают в себя подпрограммы, программы, библиотеки, объекты, классы, компоненты, структуры данных и т. д., которые выполняют определенные задачи или реализуют определенные абстрактные типы данных. Функциональные возможности программных модулей могут быть объединены или разделены между программными модулями при необходимости в различных вариантах осуществления. Исполняемые компьютером команды для программных модулей могут выполняться внутри локальной или распределенной вычислительной среды.

[00160] Любой из описанных способов может быть реализован в виде исполняемых компьютером команд или компьютерного программного продукта, хранящегося на одном или более машиночитаемых носителях (например, энергонезависимых машиночитаемых носителях, таких как один или более дисков на оптических носителях, таких как DVD или компакт-диск, компоненты энергозависимого запоминающего устройства (такие как DRAM или SRAM, или компоненты энергонезависимого запоминающего устройства, такие как жесткие диски) и исполняемые на компьютере (например, на любом имеющемся в продаже компьютере, включая смартфоны или другие мобильные устройства, включающие в себя вычислительное аппаратное обеспечение). Машиночитаемые носители не включают в себя распространяемые сигналы. Любые исполняемые компьютером команды для реализации описанных способов, а также любые данные, созданные и применяемые во время реализации описанных вариантов осуществления, могут храниться на одном или более машиночитаемых носителях (например, энергонезависимых машиночитаемых носителях).

[00161] Используемый компьютером или машиночитаемый носитель может представлять собой, например, помимо прочего, электронную, магнитную, оптическую, электромагнитную, инфракрасную или полупроводниковую систему, аппарат, устройство или среду распространения. Более конкретные примеры (неполный перечень) машиночитаемого носителя включают в себя следующее: электрическое соединение, имеющее один или более проводов, портативная компьютерная дискета, жесткий диск, оперативное запоминающее устройство (ОЗУ), постоянное запоминающее устройство

(ПЗУ), стираемое программируемое постоянное запоминающее устройство (СППЗУ или флэш-память), оптоволоконно, портативный компакт-диск - постоянное запоминающее устройство (CD-ROM), оптическое запоминающее устройство, средства передачи, такие как те, которые поддерживают Интернет или интранет, или магнитное запоминающее устройство. Следует отметить, что используемый компьютером или машиночитаемый носитель может даже представлять собой бумагу или другой приемлемый носитель, на котором печатается программа, поскольку программа может быть записана в электронном виде, например, путем оптического сканирования бумаги или другого носителя, далее может компилироваться, интерпретироваться или иным образом обрабатываться приемлемым образом, если необходимо, а затем сохраняться в запоминающем устройстве компьютера. В контексте настоящего документа используемый компьютером или машиночитаемый носитель может представлять собой любой носитель, который может содержать, хранить, передавать, распространять или переносить программу для применения или в связи с системой, аппаратом или устройством выполнения команд. Используемый компьютером носитель может включать в себя распространяемый сигнал данных с включенным в него используемым компьютером программным кодом либо в основной полосе частот, либо как часть несущей волны. Используемый компьютером программный код может передаваться с применением любого подходящего носителя, включая, помимо прочего, беспроводную связь, проводную связь, оптоволоконный кабель, радиочастоту и т. д.

[00162] Исполняемые компьютером команды могут быть частью, например, специального программного приложения или программного приложения, доступ к которому или загрузка которого осуществляется через веб-браузер или другое программное приложение (например, удаленное вычислительное приложение). Такое программное обеспечение может выполняться, например, на одном локальном компьютере (например, на любом приемлемом доступном в продаже компьютере) или в сетевой среде (например, через Интернет, глобальную вычислительную сеть, локальную вычислительную сеть, сеть «клиент-сервер» (например, сеть облачных вычислений) или другую подобную сеть, использующую один или более сетевых компьютеров.

[00163] Для ясности описаны только определенные выбранные аспекты программных вариантов реализации. Другие подробности, хорошо известные в данной области, опущены. Например, следует понимать, что описанная технология не ограничивается каким-либо конкретным компьютерным языком или программой. Например, описанная технология может быть реализована с помощью программного обеспечения, написанного на C++, Java, Perl, JavaScript, Adobe Flash, Python или любом другом приемлемом языке программирования. Аналогично, описанная технология не ограничивается каким-либо конкретным компьютером или типом аппаратного обеспечения.

[00164] Кроме того, примеры осуществления могут быть реализованы с помощью аппаратного обеспечения, программного обеспечения, встроенного программного

обеспечения, промежуточного программного обеспечения, микрокодов, языков описания аппаратного обеспечения или любой их комбинации. При реализации в программном обеспечении, встроенном программном обеспечении, промежуточном программном обеспечении или микрокоде программный код или сегменты кода для выполнения необходимых задач могут храниться на машиночитаемом или читаемом на компьютере носителе. Сегмент кода может представлять собой процедуру, функцию, подпрограмму, программу, стандартную программу, стандартную подпрограмму, модуль, программный код, программный пакет, класс или любую комбинацию команд, структуры данных, операторы программы и т. п.

[00165] В вариантах осуществления любой из вариантов осуществления, основанных на программном обеспечении (включая, например, исполняемые компьютером команды для обеспечения выполнения компьютером любого из описанных способов), может быть загружен, выгружен или удаленно доступен через приемлемые средства связи. Такие приемлемые средства связи включают в себя, например, Интернет, Всемирную сеть, интрасеть, кабель (включая оптоволоконный кабель), магнитную связь, электромагнитную связь (включая радиочастотную, микроволновую и инфракрасную связь), электронную связь или другие подобные средства связи.

[00166] В различных вариантах осуществления изделие может быть использовано для реализации одного или более способов, как описано в настоящем документе. Изделие может включать в себя машиночитаемый энергонезависимый носитель данных и носитель данных. Носитель данных может содержать команды по программированию, выполненные с возможностью обеспечения выполнения аппаратом на практике некоторых или всех аспектов описанного способа с помощью вычислительного устройства в соответствии с вариантами осуществления по настоящему описанию. Носитель данных может представлять собой широкий диапазон постоянных носителей данных, известных в данной области, включая, помимо прочего, флэш-память, оптические диски или магнитные диски. Команды по программированию, в частности, могут позволить аппарату, в ответ на их выполнение аппаратом, осуществлять различные операции, описанные в настоящем документе. Например, носитель данных может содержать команды по программированию, выполненные с возможностью обеспечения применения аппаратом на практике некоторых или всех аспектов описанного способа, например способа контроля характеристик колоночной хроматографии, включая создание модели OPLS, в соответствии с вариантами осуществления по настоящему описанию.

[00167] Хотя в настоящем документе описаны различные иллюстративные способы, аппараты, системы и изделия, объем покрытия настоящего описания не ограничивается ими. Напротив, настоящее описание охватывает все способы, аппарат и изделия, справедливо подпадающие под объем прилагаемой формулы изобретения, либо буквально, либо согласно доктрине эквивалентов. Например, хотя выше описаны иллюстративные системы, включающие в себя, среди прочих компонентов, программное обеспечение или встроенное программное обеспечение, выполняемое на аппаратном

обеспечении, следует отметить, что такие системы являются лишь иллюстративными и не должны рассматриваться как ограничивающие. В частности, предполагается, что любой из описанных компонентов аппаратного обеспечения, программного обеспечения и/или встроенного программного обеспечения может быть реализован исключительно в аппаратном обеспечении, исключительно в программном обеспечении, исключительно во встроенном программном обеспечении или в определенной комбинации аппаратного обеспечения, программного обеспечения и/или встроенного программного обеспечения.

ПРИМЕРЫ

[00168] Следующие примеры приведены для того, чтобы предоставить специалистам в данной области полное раскрытие и описание того, как получать и применять способы по настоящему изобретению, и не предназначены для ограничения объема того, что авторы настоящего изобретения считают созданным ими изобретением. Были предприняты усилия для обеспечения точности в отношении используемых чисел (например, количеств, температуры и т. д.), но следует учитывать возможность некоторых экспериментальных погрешностей и отклонений. Если не указано иное, части представляют собой части по массе, молекулярная масса представляет собой среднюю молекулярную массу, температура приведена в градусах Цельсия, комнатная температура составляет около 25 °С, а давление равно атмосферному или близко к нему.

Пример 1. Тематическое исследование: УФ-модель OPLS

[00169] УФ-модель OPLS была сгенерирована с применением способов, описанных в настоящем документе. Как проиллюстрировано на Фиг. 32-1 и Фиг. 32-2, сгенерированная модель OPLS была способна дифференцировать хорошие/удовлетворительные партии (кружки слева) и плохие/неудовлетворительные партии (кружки справа). Сгенерированная модель позволила выявить ухудшить состояние колонки до того, как это стало явно очевидным из наложения статистических данных или анализа перехода. На Фиг. 33 представлен график, сгенерированный с применением модели OPLS по заполнению колонки 1122600028. Модель была способна обнаружить ухудшение состояния колонки до того, как это было очевидно посредством визуального наблюдения (данные о колонке не были в обучающем наборе). На Фиг. 34 представлен график, сгенерированный из заполнения колонки 1122600018. Состояние данной колонки не ухудшилось, поскольку модель показала, что колонка осталась в хорошем/удовлетворительном диапазоне (левая сторона). Колонка оставалась устойчивой на протяжении всего срока службы, что свидетельствует о том, что модель работает.

Пример 2. Расширение применения моделей OPLS - процедурно сгенерированные хроматограммы (например, хроматограммы, сгенерированные искусственно, а не на основе фактических экспериментальных данных)

[00170] Как показывает пример, могут быть созданы УФ-модели OPLS, которые являются удовлетворительными, без применения партий, в которых наблюдали ухудшение состояния колонки. Ожидается, что для проточных колонок исходный УФ-отрыв в начале сбора будет расширяться по мере разрушения слоя колонки. Отрыв может

быть смоделирован посредством уравнения роста по логистической кривой. Половина данных была выбрана и дополнена процедурно сгенерированными кривыми, что позволило откалибровать колонку. На Фиг. 35 иллюстрированы сгенерированные кривые, синим цветом показаны реальные данные, а зеленым цветом - информация, введенная на основе прогноза.

[00171] На Фиг. 36 показаны результаты модели OPLS, сконструированной с применением выборочного набора реальных приемлемых статистических данных и процедурно сгенерированных кривых, представляющих ухудшение состояния (например, кривые хроматограммы, сгенерированные не на основе фактических данных), в которых, чтобы сделать модель более консервативной, верхний предел был установлен как максимальное значение для применяемых приемлемых партий. Более высокие значения указывали на большее ухудшение состояния колонки (не показаны процедурно сгенерированные разрушения).

[00172] На Фиг. 37 показана модель, применяемая для прогнозирования характеристик заполнения колонки 1194400001, в которой модель была способна прогнозировать ухудшение состояния приблизительно 10 партий перед разрушением колонки. Этот пример указывает на то, что моделирование OPLS может быть применено к УФ-сигнализации хроматографии для обнаружения незначительных, но важных изменений в характеристиках колонки. Кроме того, вместо комплексного обучающего набора с описанными моделями можно применять процедурно сгенерированные данные, которые по-прежнему позволяют моделировать ухудшение состояния колонки.

Пример 3. Прогнозирование модели OPLS

[00173] Этот пример предоставляет иллюстративные данные прогнозирования модели OPLS, сгенерированные описанными способами. На Фиг. 38А и Фиг. 38В проиллюстрировано прогнозирование модели для заполнения колонки 1122600011. ТА явно не демонстрировала развитие ухудшения состояния. Бифуркация присутствовала в следующей партии после того, как первая партия превышала $x=4$. Операция предшествовала рекомендациям по перепаковке PS. Была сконструирована модель OPLS для хроматографии гидрофобного взаимодействия (HIC) с применением хроматограмм, сгенерированных SME. Была разработана процедура для создания кривых на основе среднеквадратичного отклонения значений для выборочного набора приемлемых кривых в 2 точках. Результаты показаны на Фиг. 39А и Фиг. 39В (Набор $x=-2,46$ в качестве верхнего предела (наивысшего значения для приемлемых партий)). На Фиг. 40 проиллюстрирована сгенерированная модель OPLS, которая была способна прогнозировать ухудшение состояния для упаковок колонок 20 и 21 без применения данных из статистических отказов. Этот пример подтверждает, что успешные модели OPLS могут быть получены с применением процедурно сгенерированных кривых, и для описанных способов, которые будут применяться для контроля ухудшения состояния колонки, не требуется комплексный обучающий набор.

[00174] В целом, описанные способы обеспечивают надежный, чувствительный способ контроля ухудшения состояния колонки, который можно применять для улучшения развития обработки белка, включая развитие обработки антитела.

[00175] Настоящее изобретение не ограничивается в объеме конкретными вариантами осуществления, описанными в настоящем документе. Действительно, различные модификации данного изобретения в дополнение к описанным в настоящем документе будут очевидны специалистам в данной области из предшествующего описания и прилагаемых фигур. Подразумевается, что такие модификации входят в объем прилагаемой формулы изобретения.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Способ контроля характеристик колоночной хроматографии, включающий:
получение одной или более хроматограмм в ультрафиолетовом спектре (УФ-хроматограмм), сгенерированных системой хроматографии во время очистки и/или разделения образца; и
анализ одной или более полученных УФ-хроматограмм с помощью модели ортогональных частичных наименьших квадратов (OPLS), таким образом обеспечивая обнаружение ухудшения состояния колонки до разрушения колонки и количественный анализ УФ-сигнала в указанных одной или более УФ-хроматограмм.
2. Способ по п. 1, дополнительно включающий создание модели OPLS.
3. Способ по п. 2, в котором создание модели OPLS включает:
выбор процесса или однократной операции модели OPLS;
сбор исходных данных для УФ-кривой одного или более колоночных циклов/партий/прогонов для выбранного процесса или однократной операции;
нормализацию и совмещение собранных исходных данных;
необязательно, генерирование искусственно созданных кривых на основании нормализованных исходных данных;
классификацию и форматирование данных для импортирования во многомерный инструмент;
импортирование классифицированных и отформатированных данных во многомерный инструмент для генерирования обучающего набора; и
генерирование модели OPLS.
4. Способ по п. 3, дополнительно включающий оптимизацию сгенерированной модели OPLS, подтверждение и тестирование оптимизированной модели OPLS.
5. Способ по п. 3 или 4, в котором процесс или однократная операция представляет собой однократную операцию хроматографии для одной молекулы.
6. Способ по п. 5, в котором процесс или однократная операция представляет собой стадию аффинной хроматографии белков для одной молекулы.
7. Способ по любому из пп. 3-6, в котором сбор исходных данных для УФ-кривой одного или более колоночных циклов/партий/прогонов для выбранного процесса или однократной операции включает сбор значений УФ-поглощения в соответствующих объемах колонки для одного или более циклов/партий/прогонов.
8. Способ по любому из пп. 3-7, в котором нормализация и совмещение собранных исходных данных включает нормализацию и совмещение значений УФ-поглощения и совмещение объемов колонки.
9. Способ по п. 8, в котором нормализованные значения УФ-поглощения включают удаление вариации различий по величине в необработанном УФ-сигнале.
10. Способ по любому из пп. 3-9, в котором, необязательно, генерирование искусственно созданных кривых на основании нормализованных исходных данных включает генерирование искусственно созданных кривых на основании нормализованных

исходных данных, когда одна или более неприемлемых УФ-хроматограмм не доступны.

11. Способ по любому из пп. 1-10, дополнительно включающий обеспечение образца в системе хроматографии перед получением одной или более хроматограмм в ультрафиолетовом спектре (УФ-хроматограмм), сгенерированных системой хроматографии во время очистки и/или разделения образца.

12. Способ по любому из пп. 1-11, в котором система хроматографии представляет собой систему жидкостной хроматографии.

13. Способ по любому из пп. 1-12, в котором образец содержит белок.

14. Способ по п. 13, в котором белок представляет собой антитело, слитый белок, рекомбинантный белок или их комбинацию.

15. Способ по п. 14, в котором антитело представляет собой моноклональное антитело.

16. Способ по п. 15, в котором моноклональное антитело имеет изотип IgG1, IgG2, IgG3, IgG4 или смешанный изотип.

17. Способ создания модели ортогональных частичных наименьших квадратов (OPLS) для УФ-контроля характеристик хроматографической колонки, включающий:

выбор процесса или однократной операции модели OPLS;

сбор исходных данных для хроматограммы в ультрафиолетовом спектре (УФ-хроматограммы) одного или более колоночных циклов/партий/прогонов для выбранного процесса или однократной операции;

нормализацию и совмещение собранных исходных данных;

необязательно, генерирование искусственно созданных кривых на основании нормализованных исходных данных;

классификацию и форматирование данных для импортирования во многомерный инструмент;

импортирование классифицированных и отформатированных данных во многомерный инструмент для генерирования обучающего набора; и

генерирование модели OPLS.

18. Способ по п. 17, дополнительно включающий оптимизацию сгенерированной модели OPLS, подтверждение модели OPLS, тестирование оптимизированной модели OPLS и применение оптимизированной модели OPLS для будущих операций для контроля характеристик хроматографической колонки.

19. Способ по п. 17 или 18, в котором процесс или однократная операция представляет собой однократную операцию хроматографии для одной молекулы.

20. Способ по п. 19, в котором процесс или однократная операция представляет собой стадию аффинной хроматографии белков для одной молекулы.

21. Способ по любому из пп. 17-20, в котором сбор исходных данных для УФ-хроматограммы одного или более колоночных циклов/партий/прогонов для выбранного процесса или однократной операции включает в себя сбор значений УФ-поглощения в соответствующих объемах колонки для одного или более циклов/партий/прогонов.

22. Способ по любому из пп. 17-20, в котором нормализация и совмещение собранных исходных данных включает нормализацию и совмещение значений УФ-поглощения и совмещение объемов колонки.

23. Способ по п. 22, в котором нормализованные значения УФ-поглощения включают удаление вариации различий по величине в необработанном УФ-сигнале.

24. Способ по любому из пп. 17-23, в котором, необязательно, генерирование искусственно созданных кривых на основании нормализованных исходных данных включает генерирование искусственно созданных кривых на основании нормализованных исходных данных, когда одна или более неприемлемых УФ-хроматограмм не доступны.

25. Способ по любому из пп. 17-24, в котором указанных одна или более УФ-хроматограмм представляют собой одну или более УФ-хроматограмм, сгенерированных системой хроматографии во время очистки и/или разделения образца.

26. Способ по п. 25, в котором система хроматографии представляет собой систему жидкостной хроматографии.

27. Способ по п. 25 или п. 26, в котором образец содержит белок.

28. Способ по п. 27, в котором белок представляет собой антитело, слитый белок, рекомбинантный белок или их комбинацию.

29. Способ по п. 28, в котором антитело представляет собой моноклональное антитело.

30. Способ по п. 15, в котором моноклональное антитело имеет изотип IgG1, IgG2, IgG3, IgG4 или смешанный изотип.

31. Энергонезависимый машиночитаемый носитель данных с хранимой на нем исполняемой программой для контроля характеристик колоночной хроматографии, причем программа предписывает микропроцессору выполнение стадий:

получение одной или более хроматограмм в ультрафиолетовом спектре (УФ-хроматограмм), сгенерированных системой хроматографии во время очистки и/или разделения образца; и

анализ одной или более полученных УФ-хроматограмм с помощью модели ортогональных частичных наименьших квадратов (OPLS), таким образом обеспечивая обнаружение ухудшения состояния колонки до разрушения колонки и количественный анализ УФ-сигнала в указанных одной или более УФ-хроматограмм.

32. Энергонезависимый машиночитаемый носитель данных по п. 31, дополнительно содержащий команды для создания модели OPLS.

33. Энергонезависимый машиночитаемый носитель данных по п. 32, в котором создание модели OPLS включает:

выбор процесса или однократной операции модели OPLS;

сбор исходных данных для УФ-кривой одного или более колоночных циклов/партий/прогонов для выбранного процесса или однократной операции;

нормализацию и совмещение собранных исходных данных;

необязательно, генерирование искусственно созданных кривых на основании

нормализованных исходных данных;

классификацию и форматирование данных для импортирования во многомерный инструмент;

импортирование классифицированных и отформатированных данных во многомерный инструмент для генерирования обучающего набора; и

генерирование модели OPLS.

34. Энергонезависимый машиночитаемый носитель данных по п. 33, дополнительно включающий оптимизацию сгенерированной модели OPLS, подтверждение и тестирование оптимизированной модели OPLS.

35. Энергонезависимый машиночитаемый носитель данных по п. 33 или 34, в котором процесс или однократная операция представляет собой однократную операцию хроматографии для одной молекулы.

36. Энергонезависимый машиночитаемый носитель данных по п. 35, в котором процесс или однократная операция представляет собой стадию аффинной хроматографии белков для одной молекулы.

37. Энергонезависимый машиночитаемый носитель данных по любому из пп. 33-36, в котором сбор исходных данных для УФ-кривой одного или более колоночных циклов/партий/прогонов для выбранного процесса или однократной операции включает сбор значений УФ-поглощения в соответствующих объемах колонки для одного или более циклов/партий/прогонов.

38. Энергонезависимый машиночитаемый носитель данных по любому из пп. 33-37, в котором нормализация и совмещение собранных исходных данных включает нормализацию и совмещение значений УФ-поглощения и совмещение объемов колонки.

39. Энергонезависимый машиночитаемый носитель данных по п. 38, в котором нормализованные значения УФ-поглощения включают удаление вариации различий по величине в необработанном УФ-сигнале.

40. Энергонезависимый машиночитаемый носитель данных по любому из пп. 33-39, в котором, необязательно, генерирование искусственно созданных кривых на основании нормализованных исходных данных включает генерирование искусственно созданных кривых на основании нормализованных исходных данных, когда одна или более неприемлемых УФ-хроматограмм не доступны.

41. Энергонезависимый машиночитаемый носитель данных по любому из пп. 31-40, дополнительно включающий обеспечение образца в системе хроматографии перед получением одной или более хроматограмм в ультрафиолетовом спектре (УФ-хроматограмм), сгенерированных системой хроматографии во время очистки и/или разделения образца.

42. Энергонезависимый машиночитаемый носитель данных по любому из пп. 31-41, в котором система хроматографии представляет собой систему жидкостной хроматографии.

43. Энергонезависимый машиночитаемый носитель данных по любому из пп. 31-

42, в котором образец содержит белок.

44. Энергонезависимый машиночитаемый носитель данных по п. 43, в котором белок представляет собой антитело, слитый белок, рекомбинантный белок или их комбинацию.

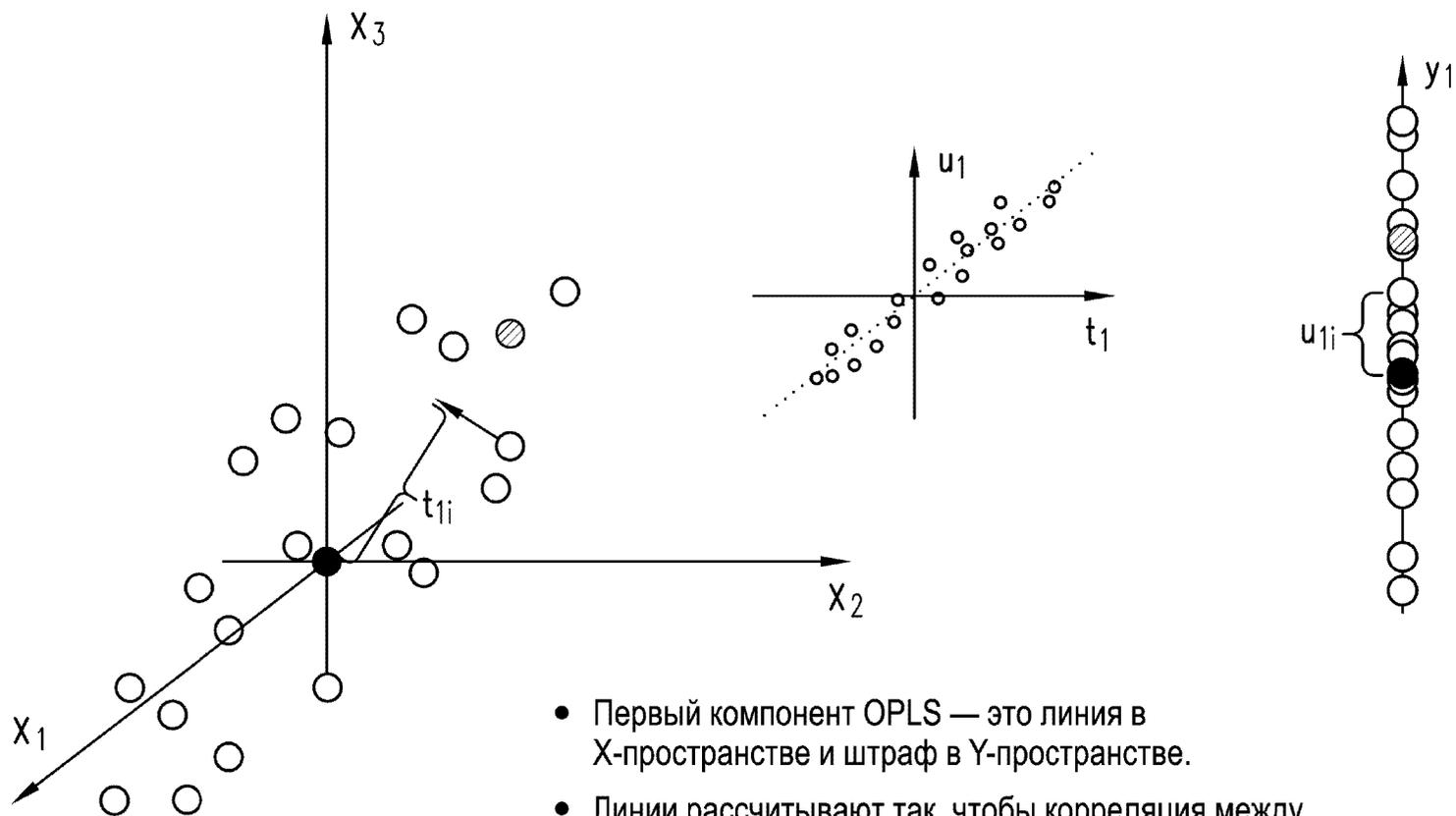
45. Энергонезависимый машиночитаемый носитель данных по п. 44, в котором антитело представляет собой моноклональное антитело.

46. Энергонезависимый машиночитаемый носитель данных по п. 45, в котором моноклональное антитело имеет изотип IgG1, IgG2, IgG3, IgG4 или смешанный изотип.

По доверенности

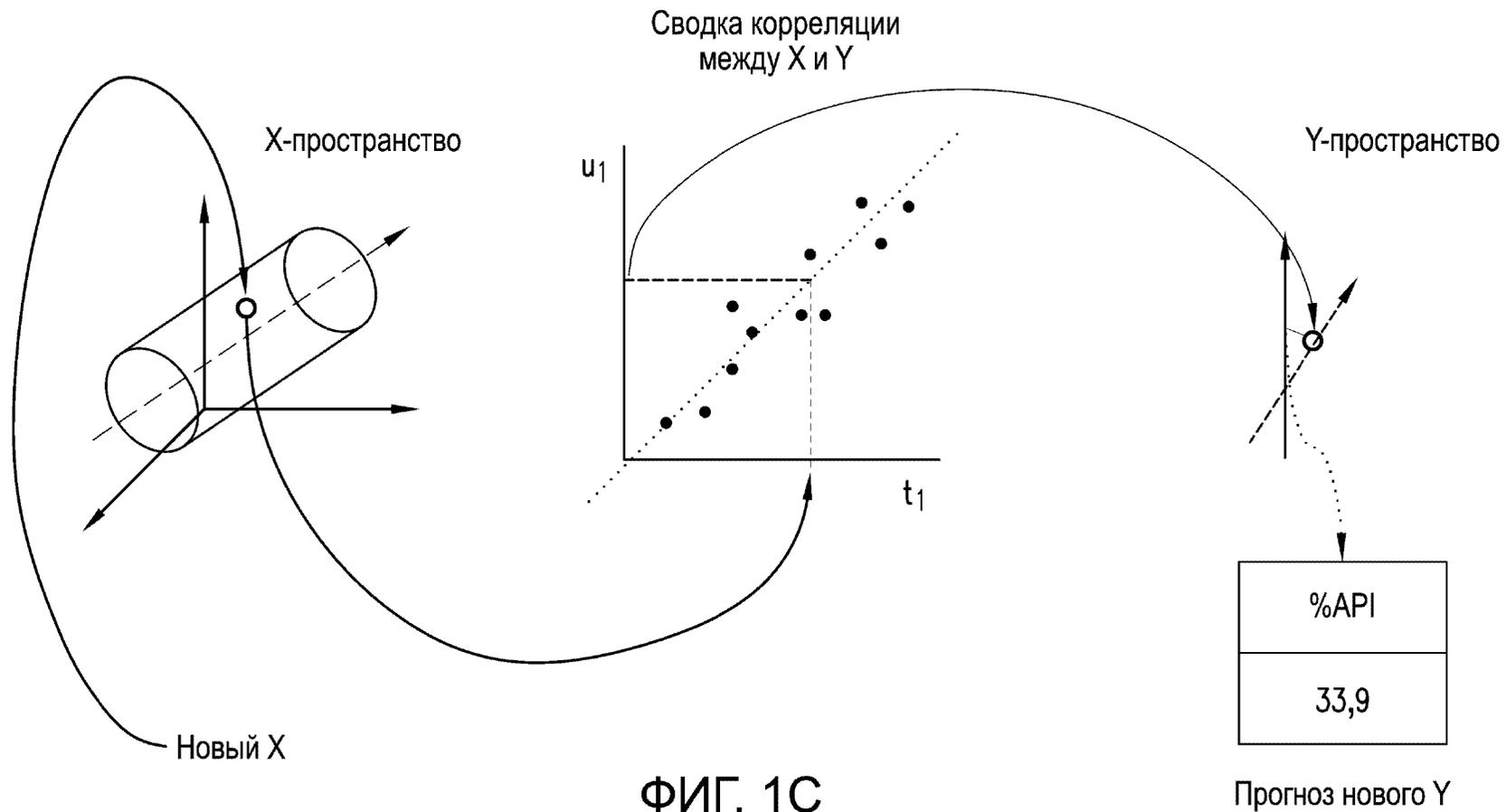


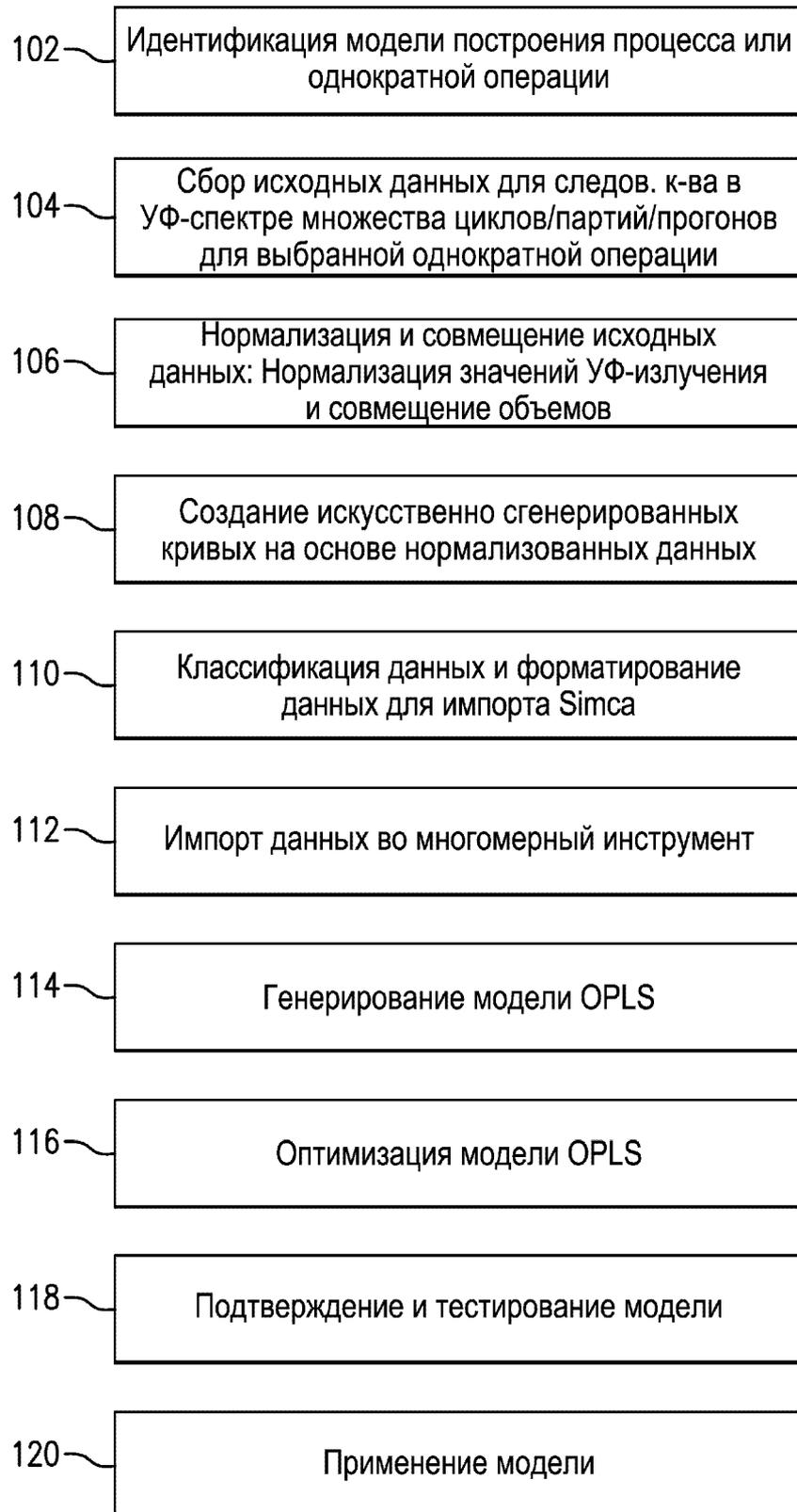
ФИГ. 1А



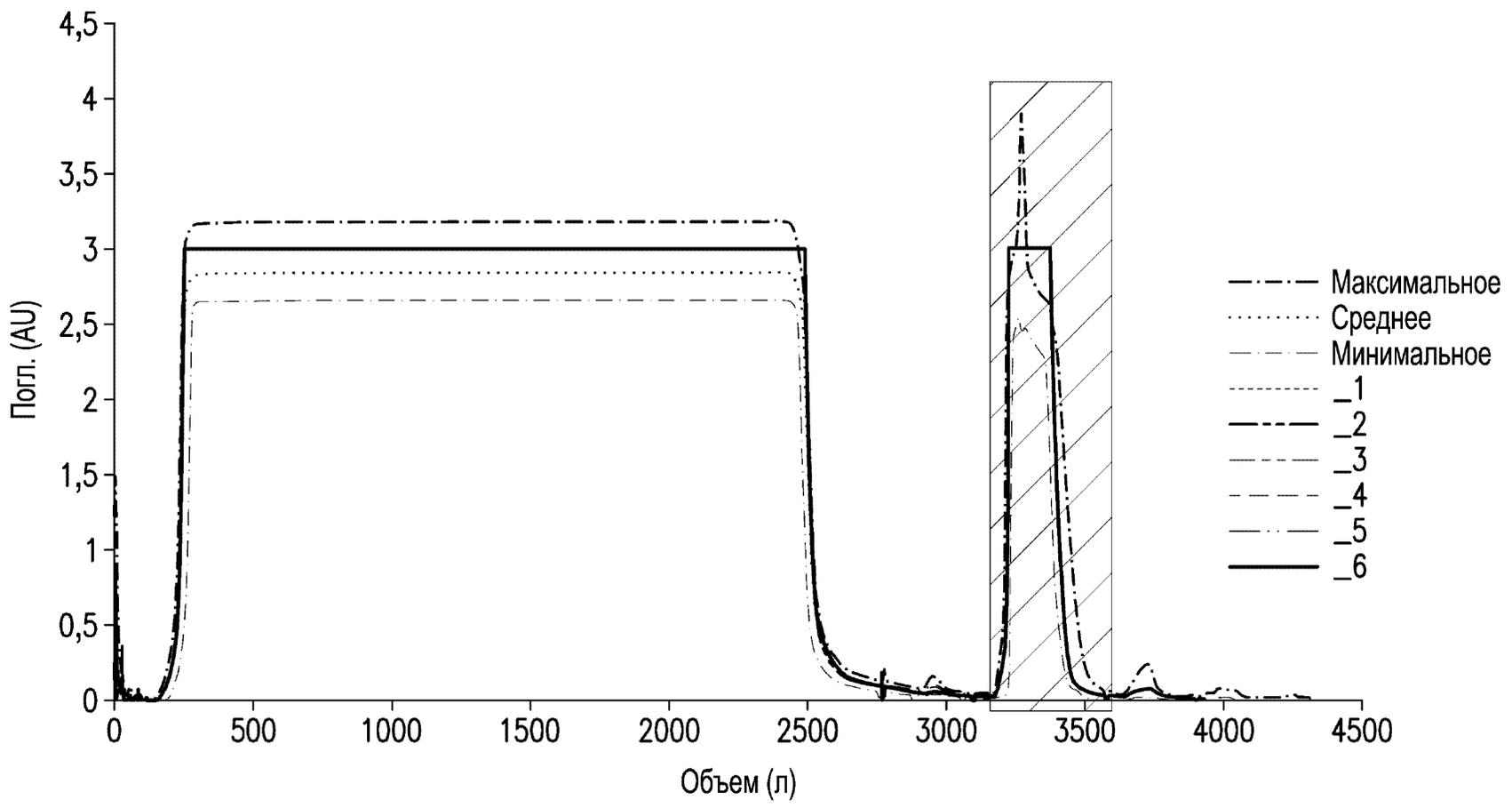
- Первый компонент OPLS — это линия в X-пространстве и штраф в Y-пространстве.
- Линии рассчитывают так, чтобы корреляция между проекциями t_1, u_1 была максимизирована.

ФИГ. 1В

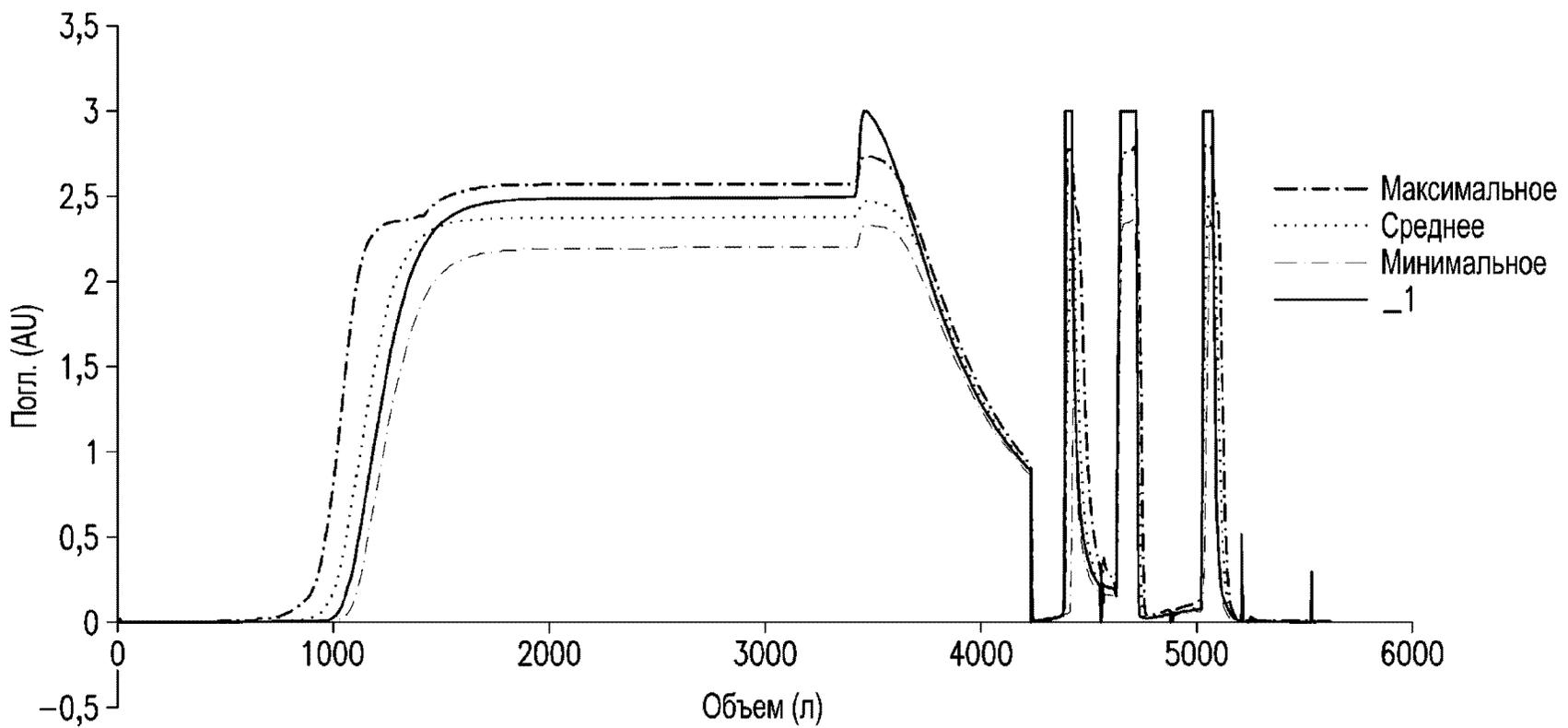




ФИГ. 1D



ФИГ. 2



ФИГ. 3

A1														
Xром. 1														
	A	B	C	D	G	H	I	J	K	L	M	N	O	P
1	Xром. 1		Xром. 1											
2	UV1:131		Веден. журнала											
3	I	AU	I	Регистрационный журнал										
4	0	0	0	Способ Прогон 14.02.2020 13:45:41 – 17:00 Способ: PA7 REGN668 ProA Загрузка версии элюирования версия 1 Результат: /ECN54013										
5	0,365351	0,0001	0	Идентификатор партии: c3a5cb72-07dc-4545-9d59-f627e5a6502f										
6	0,730702	8,73E-05	0	Базовый CV, Vc=157,000 {}, Любой										
7	1,096053	1,00E-04	0	Настройки фазового способа										
8	1,461404	8,38E-05	0	База SameAsMain										
9	1,826755	1,11E-05	0	Конечная фаза										
10	2,192106	1,00E-04	0	Определяемая пользователем фаза										
11	2,557457	7,18E-04	0	База SameAsMain										
12	2,922808	0,001177	0	Нажать_PIA111_Оповещение 4,00 {диагр.} 0,00 {диагр.} 4,00 {диагр.} 0,00 {диагр.} 0,00 {диагр.} включено										
13	3,288159	7,86E-05	0	Нажать_PIA112_Оповещение 3,00 {диагр.} 0,00 {диагр.} 3,00 {диагр.} 0,00 {диагр.} 0,00 {диагр.} включено										
14	3,65351	5,63E-05	0	AirTrapLevelControl отключен 10,0 {сек}										
15	4,018861	0	0	Блок Set_Up										
16	4,384212	0,0001	0	База SameAsMain										
17	4,749563	1,28E-07	0	Сообщение Убедитесь, что уравновешивание колонки выполнено"										
18	5,114914	6,67E-05	0	Сообщение Убедитесь, что колонка находится в режиме обхода, а MPTV и MPBV закрыты"										
19	5,480265	0,00E+00	0	Сообщение Убедитесь, что установлен соответствующий фильтр"										
20	5,845616	7,41E-05	0	Сообщение Убедитесь, что собранный материал подсоединен к ВХОДУ A1 / датчику воздуха и примирован"										
21	6,210967	0,0001	0	Сообщение Убедитесь, что Буфер 1 (основной номер элемента 10416) подсоединен к входу A2."										
22	6,576318	0,00E+00	0	Сообщение Убедитесь, что Буфер 2 (основной номер элемента 10487) подсоединен к входу B1."										
23	6,941669	0,0001	0	Сообщение Убедитесь, что Буфер 3 (основной номер элемента 10488) подсоединен к входу B2."										
24	7,30702	1,14E-05	0	Сообщение Убедитесь, что Буфер 4 (основной номер элемента 10415) подсоединен к входу B3."										
25	7,672371	1,93E-05	0	Сообщение Убедитесь, что емкость резерва установлена и подсоединен к выходу 1."										
26	8,037722	0,0001	0	Сообщение Убедитесь, что Выход 2 подсоединен к сливу"										
27	8,403073	5,17E-05	0	Сообщение Нажмите "Продолжить", когда закончите."										
28	8,768424	1,00E-04	0	Пауза Бесконечно мин. 14.02.2020 13:45:41 -17:00										
29	9,133775	8,08E-06	0	Сообщение Убедитесь, что уравновешивание колонки выполнено" подтверждено REGENERON/alissa.reyes (вручную)"										
<input type="checkbox"/> Данные 05 <input type="button" value="⊕"/>														

7/67

ФИГ. 4

Каждый столбец соответствует значениям УФ-излучения для определенной точки CV в анализе Пр.: 0.1 CV в блок элюирования

Каждая строка соответствует значениям УФ-излучения для определенного прогона/цикла Пр.: _6

		G	H	I	J	K	L	M	N	O
1	Партия_сеуеле	0	0,02	0,04	0,06	0,08	0,1	0,12	0,14	0,16
2	_1	-1,02032	-0,99778	-0,99705	-0,99707	-0,99778	-0,99781	-0,99778	-0,99842	-0,99852
3	_2	-1,01667	-0,99781	-0,99703	-0,99771	-0,99771	-0,99771	-0,99771	-0,99782	-0,99852
4	_3	-1,01593	-0,99778	-0,99698	-0,99778	-0,99778	-0,99778	-0,99778	-0,99852	-0,99852
5	_4	-1,01742	-0,99705	-0,99704	-0,99703	-0,99778	-0,99778	-0,99771	-0,99778	-0,99845
6	_5	-1,01593	-0,99776	-0,99779	-0,99779	-0,99779	-0,99779	-0,99846	-0,99846	-0,99852
7	_6	-1,01525	-0,99785	-0,99711	-0,99711	-0,99778	-0,99778	-0,99785	-0,99785	-0,99859
8	_1	-1,02035	-0,99784	-0,99707	-0,99784	-0,99778	-0,99778	-0,99854	-0,99852	-0,99853
9	_2	-1,01742	-0,99778	-0,99704	-0,99778	-0,99773	-0,99773	-0,99778	-0,99845	-0,99845
10	_3	-1,01742	-0,99704	-0,99704	-0,99704	-0,99704	-0,99771	-0,99775	-0,99771	-0,99775
11	_4	-1,01667	-0,99697	-0,99701	-0,99704	-0,99701	-0,99698	-0,99771	-0,99771	-0,99771
12	_5	-1,01593	-0,99706	-0,99627	-0,99698	-0,99698	-0,99698	-0,99698	-0,99779	-0,99778
13	_6	-1,01667	-0,99704	-0,99704	-0,99704	-0,99779	-0,99778	-0,99771	-0,99775	-0,99852
14										
15										
16										
17										
18										
19										
20										
21										
22										
23										
24										
25										
26										
27										
28										
29										

Все значения УФ-излучения нормализованы до -1 до 1

Продолж. На ФИГ. 5-2

8/67

ФИГ. 5-1

Кривая для подгонки
 посредством
 математической модели

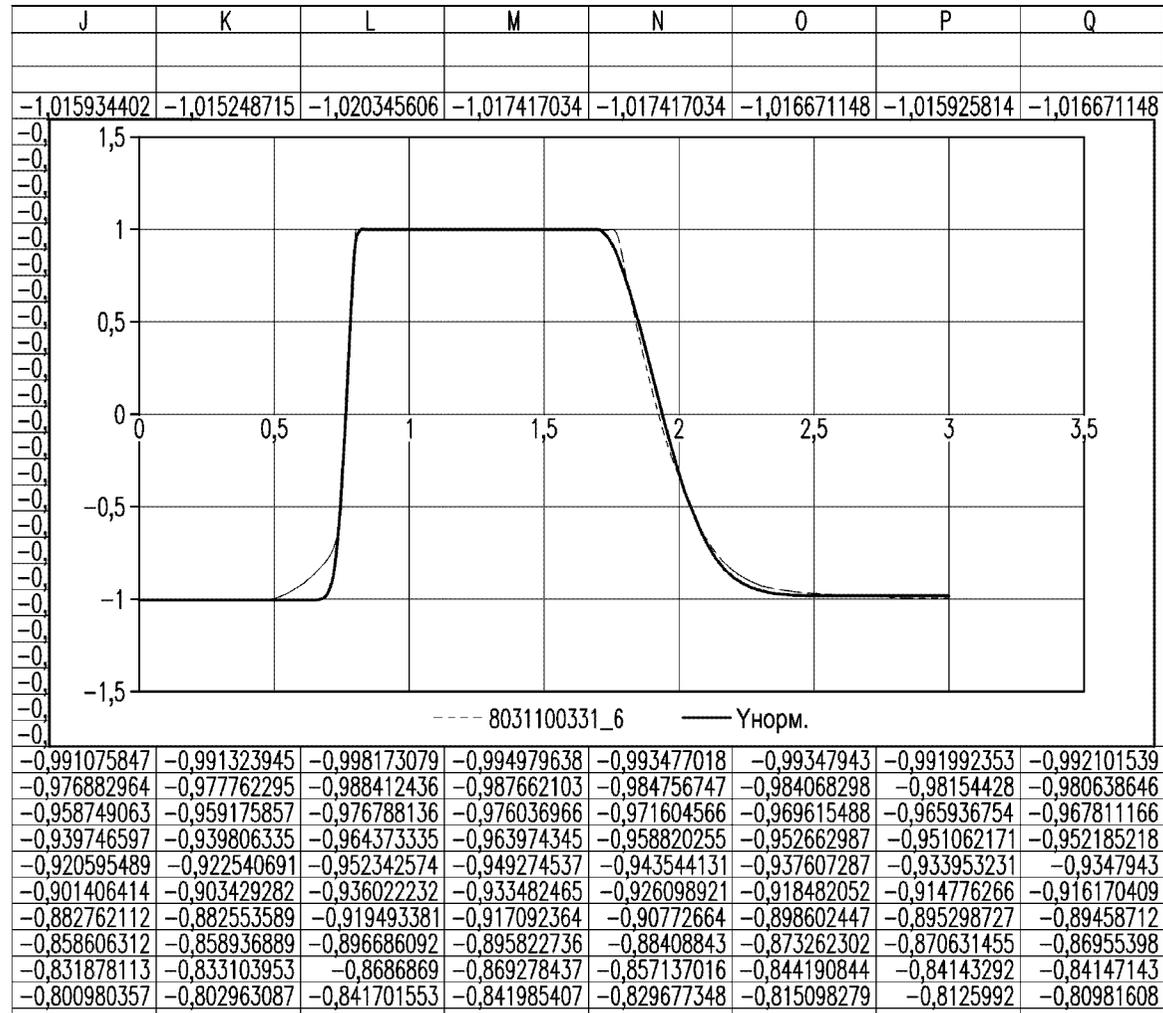
B1		X ✓ f _x				
	A	E	F	G	H	I
1						
2	Нормализ.					
3	0	-1,016671148	-1,020315753	-1,016671148	-1,015926146	-1,017417034
4	0,02	-0,997042216	-0,997781731	-0,997811502	-0,997782811	-0,997045385
5	0,04	-0,997042216	-0,99705085	-0,997026311	-0,99697644	-0,997041122
6	0,06	-0,997042216	-0,997074679	-0,997714439	-0,997782811	-0,997033556
7	0,08	-0,997788236	-0,997781731	-0,997714439	-0,997782811	-0,997780841
8	0,1	-0,997799287	-0,997807756	-0,997714439	-0,997782811	-0,997780841
9	0,12	-0,997714439	-0,997781731	-0,997714439	-0,997782811	-0,997713594
10	0,14	-0,99775084	-0,998422565	-0,997819696	-0,998521983	-0,997780841
11	0,16	-0,998521108	-0,998522798	-0,998521108	-0,998521983	-0,998453314
12	0,18	-0,998578818	-0,998522798	-0,998569313	-0,998521983	-0,998453314
13	0,2	-0,998521108	-0,998555411	-0,998521108	-0,998521983	-0,998453314
14	0,22	-0,998521108	-0,998522798	-0,998521108	-0,999261156	-0,998453314
15	0,24	-0,999260554	-0,999263865	-0,999211367	-0,999261156	-0,998570747
16	0,26	-0,999260554	-0,999252007	-0,999207926	-0,999261156	-0,99926028
17	0,28	-0,999260554	-0,999263865	-0,999260554	-0,999210678	-0,999128745
18	0,3	-0,999260554	-0,999263865	-0,999212187	-0,999241326	-0,99926028
19	0,32	-0,999260554	-0,999263865	-0,999244777	-0,999261156	-0,99926028
20	0,34	-0,999193822	-0,999937562	-0,999260554	-0,999933132	-0,999179718
21	0,36	-0,999980768	-0,999937562	-0,999204806	-0,999933132	-0,999932753
22	0,38	-1	-0,999937562	-1	-0,999933132	-0,999985401
23	0,4	-0,999948192	-0,999937562	-0,999986226	-0,999933132	-0,999979335
24	0,42	-1	-0,999937562	-1	-0,999933132	-1
25	0,44	-0,999932778	-0,999938215	-0,999993439	-0,999933132	-0,999969329
26	0,46	-1	-0,999937562	-0,999958685	-0,999933132	-0,99999885
27	0,48	-0,998521108	-0,999216423	-0,999260554	-0,998521983	-0,998112969
28	0,5	-0,992101539	-0,994143766	-0,994860769	-0,992003822	-0,989787297
29	0,52	-0,980638646	-0,983354754	-0,984807744	-0,980533693	-0,975663555
30	0,54	-0,967811166	-0,970614809	-0,971311004	-0,966940928	-0,959192408
31	0,56	-0,952185218	-0,957808257	-0,956500222	-0,952173711	-0,941361756
32	0,58	-0,9347943	-0,945767724	-0,940809895	-0,938682307	-0,924289768
33	0,6	-0,916170409	-0,931018336	-0,924513848	-0,922575796	-0,906550284
34	0,62	-0,89458712	-0,914367169	-0,906074415	-0,904114482	-0,888780606
35	0,64	-0,86955398	-0,894298806	-0,884811172	-0,88284962	-0,866013113
36	0,66	-0,84147143	-0,871514097	-0,859812987	-0,858330582	-0,840916042
37	0,68	-0,80981608	-0,845002023	-0,831952272	-0,831057655	-0,811299687

Продолж.
 на ФИГ. 6-2

10/67

ФИГ. 6-1

Нормализованные кривые импортированы



Продолж.
От ФИГ. 6-1

Продолж. От
ФИГ. 6-3

11/67

ФИГ. 6-2

Значение RMSE

R	S	T	U	V	W	
	Описание	Начало:	$\Sigma((p - o)^2)$:	0,25146	RMSE:	0,040
	высота пикового отрыва	hm1	0,85			Исходн
	Местополож. отрыва	tm1	0,81			7,68
	крутизна отрыва	s1	0,04			1,69
						2,90
		Конец:				3,87
	Высота пиков. падения	hm2	0,84			4,03
	Местопол. падения	tm2	1,69			3,26
	крутизна падения	s2	0,21			2,06
						1,0
		Границы:				3,88
	Точка, где eqn1 -> eqn pmstart		0,81			1,1
	Точка, где eqn2 -> eqn pmend		1,69			2,6
						4,85
	ht 1/3 eqns	равномерн. h	0,85			6,83
						7,49
		Макс.:				6,40
	Высота пиков. макс.	hm3	0			4,26
	Местопол. пиков. макс.	tm3	0,92			2,20
	Ширина пиков. макс.	s3	0,01			8,90
						2,79
		Фактор затягивания.				6,84
	увеличивает затягивание	затягивание	0,01			1,30
						1,93
	Ключ					2,23
	Параметр должен оставаться постоянным					2,01
	Параметр должен быть изменен					1,4
						7,70
						3,27
						1,08
						2,79
						5,62
						8,79
						1,07
						0,000
						0,000
						0,004

Продолж.
От ФИГ. 6-2

12/67

ФИГ. 6-3

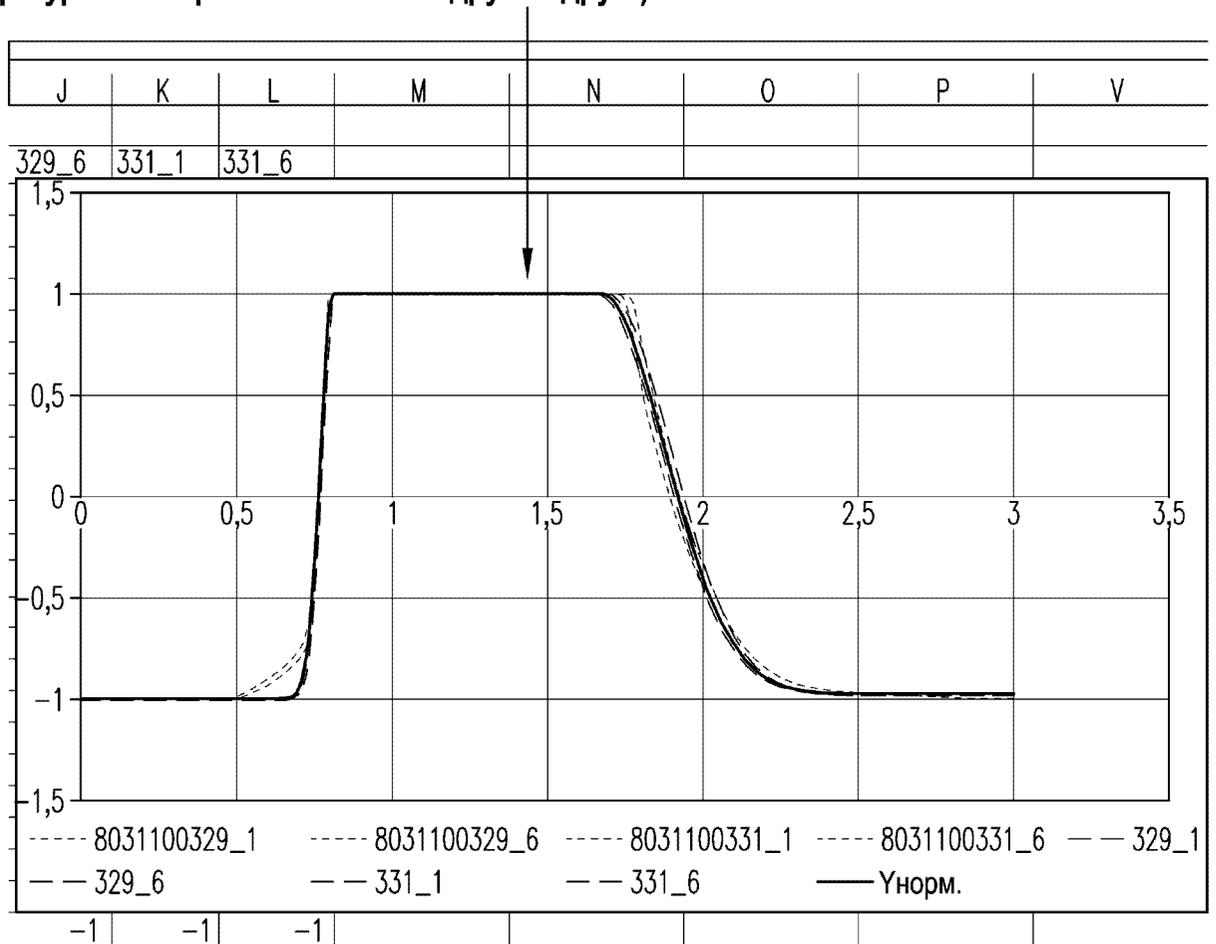
Подгонка модели к существующим хроматограммам и копирование в пространство справа с помощью "CTL-SHIFT-C"

R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD
	Описание	Начало:	$\sum((p - o)^2)$:	0,25146	RMSE:	0,04080844						
	высота пикового отрыва	hm1	0,85			Исходн.	Унорм.	329_1	329_6	331_1	331_6	
	Местополож. отрыва	tm1	0,81			7,6821E-90	-1	-1	-1	-1	-1	-1
	крутизна отрыва	s1	0,04			1,6921E-85	-1	-1	-1	-1	-1	-1
						2,9026E-81	-1	-1	-1	-1	-1	-1
		Конец:				3,8779E-77	-1	-1	-1	-1	-1	-1
	Высота пиков. падения	hm2	0,84			4,0347E-73	-1	-1	-1	-1	-1	-1
	Местопол. падения	tm2	1,69			3,2694E-69	-1	-1	-1	-1	-1	-1
	Крутизна падения	s2	0,21			2,0632E-65	-1	-1	-1	-1	-1	-1
						1,014E-61	-1	-1	-1	-1	-1	-1
		Границы:				3,8813E-58	-1	-1	-1	-1	-1	-1
	Точка, где eqn1 -> eqn pmstart		0,81			1,157E-54	-1	-1	-1	-1	-1	-1
	Точка, где eqn2 -> eqn pmend		1,69			2,686E-51	-1	-1	-1	-1	-1	-1
						4,8565E-48	-1	-1	-1	-1	-1	-1
	ht 1/3 eqns	равномерн. h	0,85			6,8384E-45	-1	-1	-1	-1	-1	-1
						7,4993E-42	-1	-1	-1	-1	-1	-1
		Макс.:				6,4048E-39	-1	-1	-1	-1	-1	-1
	Высота пиков. макс.	hm3	0			4,2601E-36	-1	-1	-1	-1	-1	-1
	Местопол. пиков. макс.	tm3	0,92			2,2068E-33	-1	-1	-1	-1	-1	-1
	Ширина пиков. макс.	s3	0,01			8,9028E-31	-1	-1	-1	-1	-1	-1
						2,7972E-28	-1	-1	-1	-1	-1	-1
		Фактор затягивания:				6,8445E-26	-1	-1	-1	-1	-1	-1
	увеличивает затягивание	затягивание	0,01			1,3043E-23	-1	-1	-1	-1	-1	-1
						1,9358E-21	-1	-1	-1	-1	-1	-1
	ключ:					2,2375E-19	-1	-1	-1	-1	-1	-1
	Параметр должен оставаться постоянным					2,0141E-17	-1	-1	-1	-1	-1	-1
	Параметр должен быть изменен					1,412E-15	-1	-1	-1	-1	-1	-1
						7,7093E-14	-1	-1	-1	-1	-1	-1
						3,2781E-12	-1	-1	-1	-1	-1	-1
						1,0855E-10	-1	-1	-1	-1	-1	-1
						2,7997E-09	-1	-1	-1	-1	-1	-1
						5,6233E-08	-1	-1	-1	-1	-1	-1
						8,7963E-07	-1	-1	-1	-1	-1	-1
						1,0716E-05	-0,99997	-1	-0,99997	-0,99997	-0,99997	-0,99997
						0,00010167	-0,99976	-0,9999	-0,99976	-0,99976	-0,99976	-0,99976
						0,00075125	-0,99823	-0,9993	-0,99823	-0,99823	-0,99823	-0,99823
						0,00432316	-0,98983	-0,9956	-0,98983	-0,98983	-0,98983	-0,98983

13/67

ФИГ. 7

Диаграмма, показывающая сгенерированные кривые. Реальные данные выделены синим, кривые, сгенерированные на предыдущей стадии, выделены зеленым, искусственная кривая, сгенерированная на основе управляемых параметров, выделена коричневым (на фигуре ниже кривые наложены друг на друга)



Продолж.
На ФИГ. 9-2

15/67

ФИГ. 9-1

Скопировать средние значения
и значения SD из вкладки
генерирования в таблицу

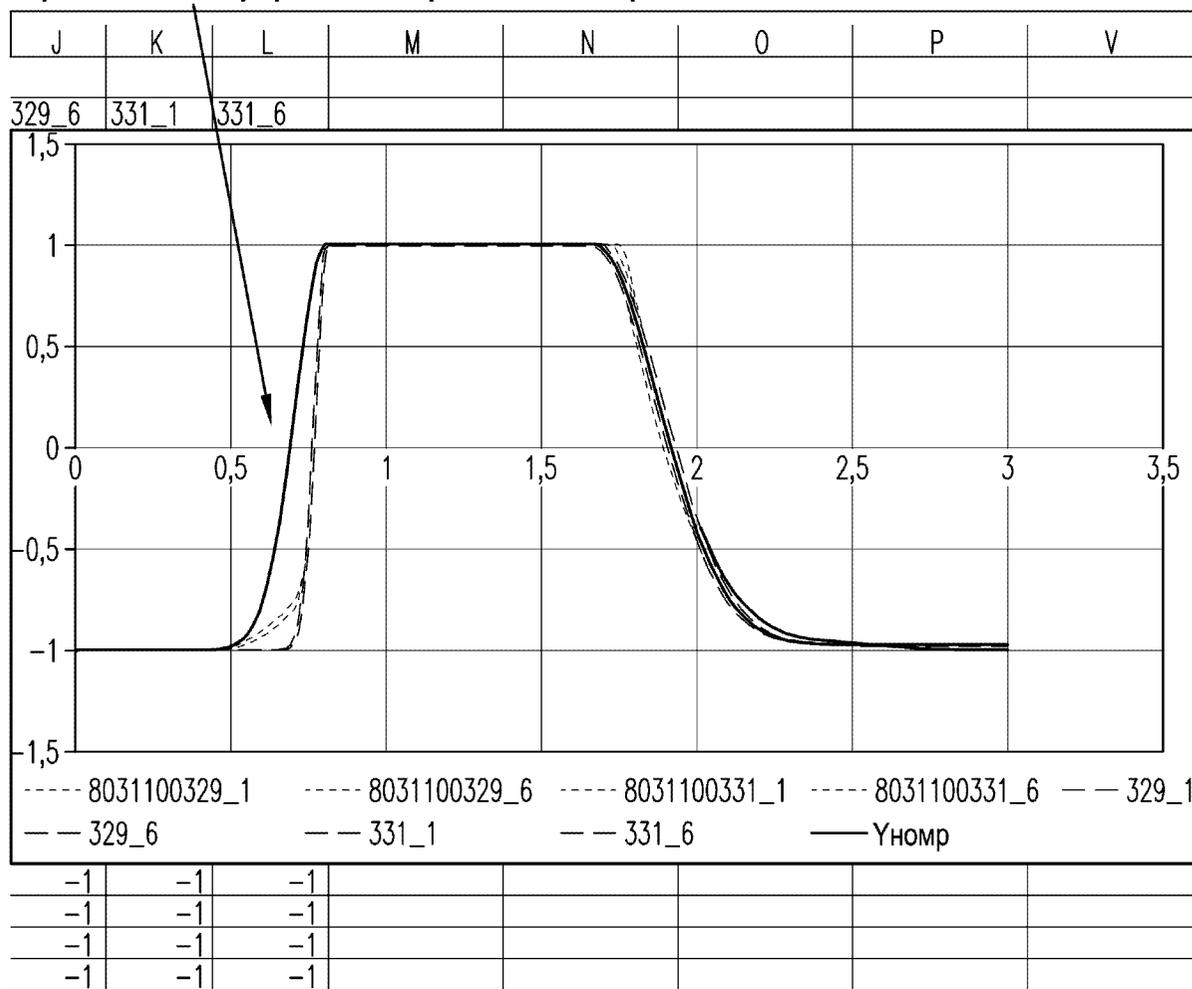
Коррекция значений SD для
получения хроматограмм, типичных
для ухудшения состояния колонки.
Значения параметров будут равны
среднему значению и указанному
количеству среднеквадр.
отклонений

Продолж. От
ФИГ. 9-1

W	X	Y	Z	AA	AB	AC	AD
				Описание	Начало:		Станд. откл.
				высота пикового отрыва	hm1	0,85	0
				Местополож. отрыва	tm1	0,8125	0
	СРЕД	Станд. откл.		крутизна отрыва	s1	0,04	0
tm1	0,8125	0,00433					
s1	0,04	0			Конец:		
tm2	1,6725	0,01299		Высота пиков. падения	hm2	0,84	0
s2	0,21	0		Местопол. падения	tm2	1,6725	0
hm3	0	0		крутизна падения	s2	0,21	0
tm3	1,025	0,139555					
s3	0,31	0,51962			Границы:		
затягивание	0,01	0		Точка, где eqn1 -> eqn pmstart	0,8125		0
	0,2548	0,05386		Точка, где eqn2 -> eqn pmend	1,6725		0
				ht 1/3 eqns	равномер	0,85	
					Макс:		
				Высота пиков. макс.	hm3	0	0
				Местопол. пиков. макс.	tm3	1,025	0
				Ширина пиков. макс.	s3	0,31	0
					Фактор затягивания:		
				увеличивает затягивание	затягивание	0,01	0
				Ключ			
				Параметр должен оставаться постоянным			

ФИГ. 9-2

Пиковый отрыв для сгенерированной кривой стал шире



Продолж. На
ФИГ. 10-2

17/67

ФИГ. 10-1

S1 Значение было вручную изменено на 0,1 (SD было равно 0, поэтому пришлось корректировать вручную)

	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG
					Описание	Начало		SD	0,276329	0,27632918	
					высота пикового отрыва	hm1	0,85	0		Исходн.	Унорм.
					Местополож. отрыва	tm1	0,8125	0		3,9292E-15	-1
		СРЕД	Станд. откл.		Крутизна отрыва	s1	0,1	0		1,9559E-14	-1
tm1	0,8125	0,00433								9,3545E-14	-1
s1	0,04	0				Конец:				4,2985E-13	-1
tm2	1,6725	0,01299			Высота пиков. падения	hm2	0,84	0		1,8978E-12	-1
s2	0,21	0			Местопол. падения	tm2	1,6725	0		8,0501E-12	-1
hm3	0	0			Крутизна падения	s2	0,21	0		3,2809E-11	-1
tm3	1,025	0,13955								1,2847E-10	-1
s3	0,31	0,51962				Границы:				4,8333E-10	-1
затягивание	0,01	0			Точка, где eqn1 -> eqn pmstart		0,8125	0		1,7471E-09	-1
	0,2548	0,05386			Точка, где eqn2 -> eqn pmend		1,6725	0		6,0674E-09	-1
										2,0246E-08	-1
					ht 1/3 eqns	равномер	0,85			6,4906E-08	-1
										1,9992E-07	-1
						Макс:				5,9166E-07	-1
					Высота пиков. макс.	hm3	0	0		1,6823E-06	-1
					Местопол. пиков. макс.	tm3	1,025	0		4,596E-06	-0,99999
					Ширина пиков. макс.	s3	0,31	0		1,2063E-05	-0,99997
										3,0422E-05	-0,99993
						Фактор затягивания:				7,3713E-05	-0,99983
					увеличивает затягивание	затягивание	0,01	0		0,0001716	-0,9996
										0,00038382	-0,9991
					Ключ					0,00082483	-0,99806
					Параметр должен оставаться постоянным					0,00170306	-0,99599
					Параметр должен быть изменен					0,00337847	-0,99205
										0,00643933	-0,98485
										0,01179203	-0,97225

Продолж. От
ФИГ. 10-1

Сохраняются нормализованные значения Уф-излучения для искусственной кривой. Скопируйте значения в столбцы справа, чтобы сохранить значения. Назовите кривую в зависимости от того, какой параметр был изменен (Пр.: s1:0.1)

18/67

ФИГ. 10-2

Классификационные группы «Хорошая» или «плохая» классификация для каждой партии Нормализованные значения УФ-излучения для каждого прогона

J19		✕ ✓ f_x -0,999999994439936								
	A	B	C	D	G	H	I	J	K	L
1	Партия_цикл	Область обработки	Заполнение колонки	Хорошая/Плохая	0	0,02	0,04	0,06	0,08	0,1
2	-1	PA7	32	Хорошая	-1,02032	-0,99778	-0,99705	-0,99707	-0,99778	-0,99781
3	-2	PA7	32	Хорошая	-1,01667	-0,99781	-0,99703	-0,99771	-0,99771	-0,99771
4	-3	PA7	32	Хорошая	-1,01593	-0,99778	-0,99698	-0,99778	-0,99778	-0,99778
5	-4	PA7	32	Хорошая	-1,01742	-0,99705	-0,99704	-0,99703	-0,99778	-0,99778
6	-5	PA7	32	Хорошая	-1,01593	-0,99776	-0,99779	-0,99779	-0,99779	-0,99779
7	-6	PA7	32	Хорошая	-1,01525	-0,99785	-0,99711	-0,99711	-0,99778	-0,99778
8	-1	PA7	32	Хорошая	-1,02035	-0,99784	-0,99707	-0,99784	-0,99778	-0,99778
9	-2	PA7	32	Хорошая	-1,01742	-0,99778	-0,99704	-0,99778	-0,99773	-0,99773
10	-3	PA7	32	Хорошая	-1,01742	-0,99704	-0,99704	-0,99704	-0,99704	-0,99771
11	-4	PA7	32	Хорошая	-1,01667	-0,99697	-0,99701	-0,99704	-0,99701	-0,99698
12	-5	PA7	32	Хорошая	-1,01593	-0,99706	-0,99627	-0,99698	-0,99698	-0,99698
13	-6	PA7	32	Хорошая	-1,01667	-0,99704	-0,99704	-0,99704	-0,99779	-0,99778
14	329_1	PGC	PGC	Хорошая	-1	-1	-1	-1	-1	-1
15	329_6	PGC	PGC	Хорошая	-1	-1	-1	-1	-1	-1
16	331_1	PGC	PGC	Хорошая	-1	-1	-1	-1	-1	-1
17	331_6	PGC	PGC	Хорошая	-1	-1	-1	-1	-1	-1
18	s1:0,1	PGC	PGC	Плохая	-3,35294	-3,35294	-3,35294	-3,35294	-3,35294	-3,35294
19	s1:0,12	PGC	PGC	Плохая	-1	-1	-1	-1	-1	-1
20	s2:0,5	PGC	PGC	Плохая	-1	-1	-1	-1	-1	-1
21	s2:0,4	PGC	PGC	Плохая	-1	-1	-1	-1	-1	-1
22	s2:0,35	PGC	PGC	Плохая	-1	-1	-1	-1	-1	-1
23	tm1:0,65	PGC	PGC	Плохая	-1	-1	-1	-1	-1	-1
24	tm1:0,68	PGC	PGC	Плохая	-1	-1	-1	-1	-1	-1
25	tm2:2	PGC	PGC	Плохая	-1	-1	-1	-1	-1	-1
26	tm2:1,75, s2:0,35	PGC	PGC	Плохая	-1	-1	-1	-1	-1	-1
27	tm1:0,78, s1:0,1	PGC	PGC	Плохая	-1	-1	-1	-1	-1	-1

Нормализованные исходные данные импортированы

Искусственно сгенерированные «хорошие» кривые

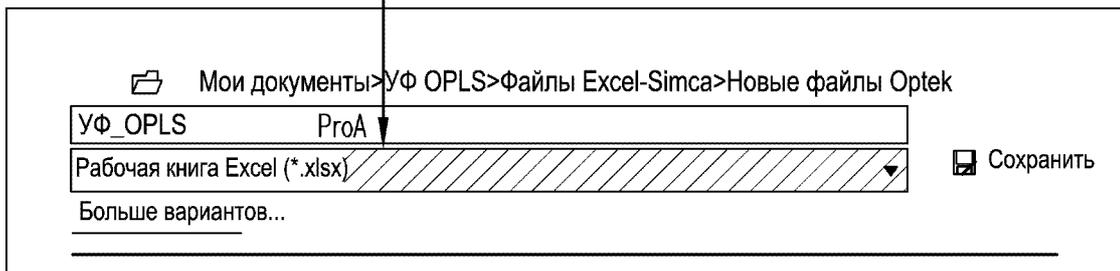
Искусственно сгенерированные «плохие» кривые

Продолж. На ФИГ. 11-2

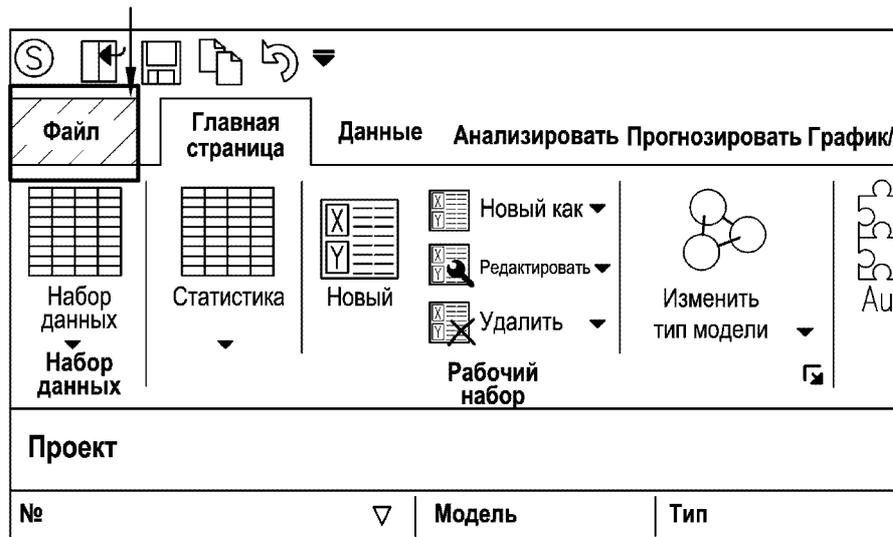
19/67

ФИГ. 11-1

1. Сохранить файл Excel как обычный файл Excel (без включения макросов)

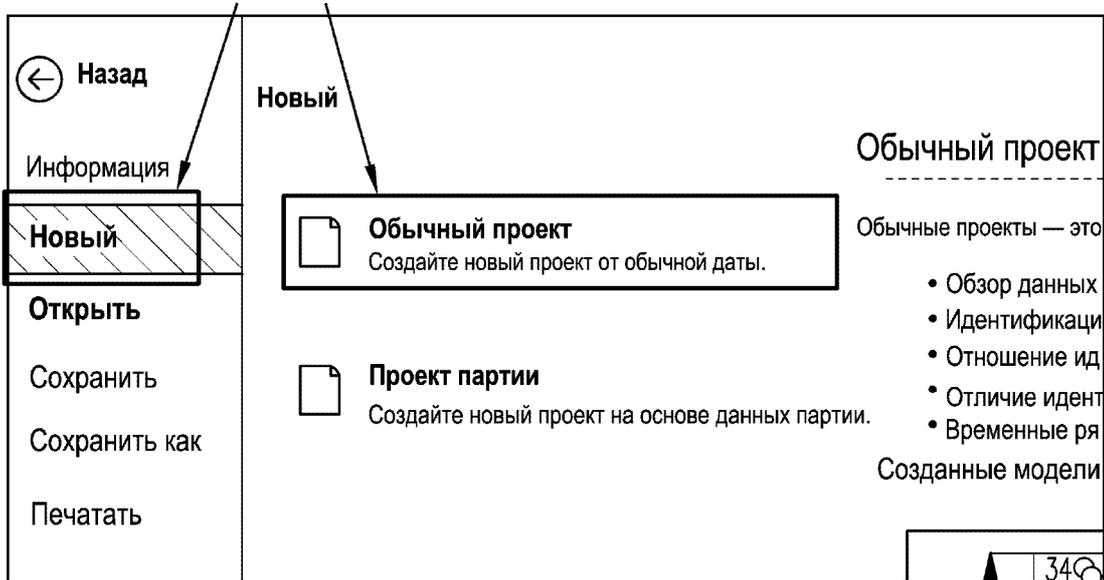


2. Открыть Simca и выбрать «Файл»

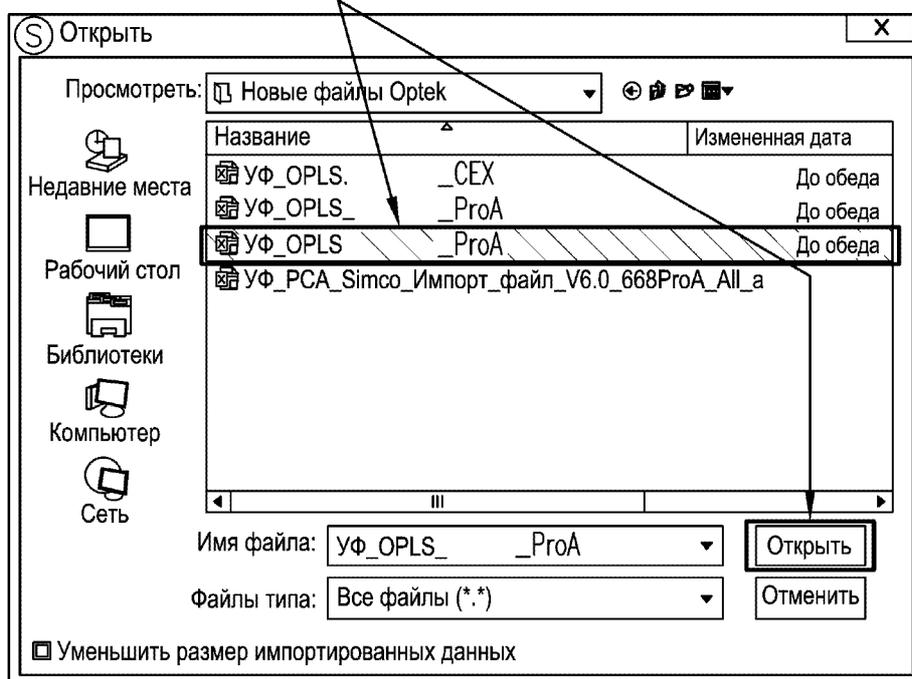


ФИГ. 12А

3. Выбрать "Новый", затем Обычный проект

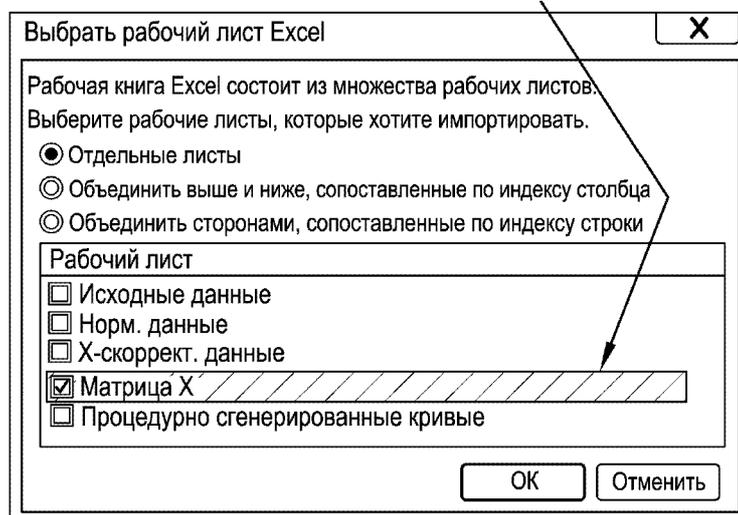


4. Выбрать файл Excel и выбрать «Открыть».



ФИГ. 12В

5. Снимите галочки со всех листов, кроме «Матрица X». Затем нажмите «ОК».



ФИГ. 13А

Импорт SIMCA – [Матрица X]

Файл Главная страница Редактировать Просмотреть

Добавить данные Первичный идентиф. переменной Вторичный идентиф. переменной Идентификатор первичного наблюдения Вторичный идентификатор наблюдения Идентификатор класса Количественные Качественные

Типы данных

Матрица X ж (+)

	Первичн. идент.	Качественные	3	Качественные	5	6	7
1	Партия_цикл	Область обработки	Заполнение колонки	Хорошая/Плохая	Бифуркация?	Сильное затягивание?	0
2	1 PA7		32	Хорошая			-1,02032
3	2 PA7		32	Хорошая			-1,01667
4	3 PA7		32	Хорошая			-1,01593
5	4 PA7		32	Хорошая			-1,01742
6	5 PA7		32	Хорошая			-1,01593
7	6 PA7		32	Хорошая			-1,01525
8	1 PA7		32	Хорошая			-1,02035
9	2 PA7		32	Хорошая			-1,01742
10	3 PA7		32	Хорошая			-1,01742
11	4 PA7		32	Хорошая			-1,01667
12	5 PA7		32	Хорошая			-1,01593
13	6 PA7		32	Хорошая			-1,01667
14	329_1	PCC	PCC	Хорошая			-1
15	329_6	PCC	PCC	Хорошая			-1
16	331_1	PCC	PCC	Хорошая			-1
17	331_6	PCC	PCC	Хорошая			-1
18	s1:0,1	PCC	PCC	Плохая			-1
19	s1:0,12	PCC	PCC	Плохая			-1
20	s2:0,5	PCC	PCC	Плохая			-1
21	s2:0,4	PCC	PCC	Плохая			-1
22	s2:0,35	PCC	PCC	Плохая			-1
23	tm1:0,65	PCC	PCC	Плохая			-1
24	tm1:0,68	PCC	PCC	Плохая			-1
25	tm2:2	PCC	PCC	Плохая			-1
26	tm2:1,75, s2:0,35	PCC	PCC	Плохая			-1
27	tm1:0,78, s1:0,1	PCC	PCC	Плохая			-1

III

вопросов

► Решение всех

Лист: Матрица X (155 переменных, 27 наблюдений, 69 (1,65%) отсутствуют)

ⓘ Нет первичн. идент. верс.

Продолж. На
ФИГ. 13В-2

6. Убедитесь, что первый столбец является перв. идент.

ФИГ. 13В-1

7. Выберите строку 1 и нажмите Первичный идентиф. переменной. Теперь строка должна быть окрашена зеленым.

Импорт SIMCA — [Матрица X]

Файл Главная страница Редактировать Просмотреть

Добавить данные Первичный идентиф. переменной Вторичный идентиф. переменной Идентификатор первичного наблюдения Вторичный идентификатор наблюдения

Идентификаторы переменных Идентификаторы наблюдения

Идентификатор класса Количественные Качественные Дата/Время Переменная X Переменная Y Исключить Включить Очистить форматирование Применить форматирование Завершить импорт

Типы данных Роли переменных Исключить Автоформатирование Импорт

Матрица X: (+)

	Первичн. идентиф.	Качественные	3	Качественные	5	6	7	8	9	10	11	12	13	14	15
Первичн. идентиф.	Партия_цикл	Область обработки	Заполнение колонки	Хорошая/Плохая	Бифуркация?	Сильное затягивание?	0	0,02	0,04	0,06	0,08	0,1	0,12	0,14	0,16
2	1	PA7	32	Хорошая			-1,02032	-0,997782	-0,997051	-0,997075	-0,997782	-0,997808	-0,997782	-0,998423	-0,998523
3	2	PA7	32	Хорошая			-1,01667	-0,997811	-0,997026	-0,997714	-0,997714	-0,997714	-0,997714	-0,99782	-0,998521
4	3	PA7	32	Хорошая			-1,01593	-0,997783	-0,996976	-0,997783	-0,997783	-0,997783	-0,997783	-0,998522	-0,998522
5	4	PA7	32	Хорошая			-1,01742	-0,997045	-0,997041	-0,997034	-0,997781	-0,997781	-0,997714	-0,997781	-0,998453
6	5	PA7	32	Хорошая			-1,01593	-0,997758	-0,997791	-0,997791	-0,997791	-0,997791	-0,998463	-0,998463	-0,998522
7	6	PA7	32	Хорошая			-1,01525	-0,99785	-0,997111	-0,997111	-0,997783	-0,9978	-0,99785	-0,99785	-0,998589
8	1	PA7	32	Хорошая			-1,02035	-0,997844	-0,997072	-0,997844	-0,997777	-0,997777	-0,99854	-0,998518	-0,998535
9	2	PA7	32	Хорошая			-1,01742	-0,997781	-0,997041	-0,997781	-0,997731	-0,997728	-0,997781	-0,998453	-0,998453
10	3	PA7	32	Хорошая			-1,01742	-0,997041	-0,997041	-0,997041	-0,997041	-0,997714	-0,997754	-0,997714	-0,997749
11	4	PA7	32	Хорошая			-1,01667	-0,996975	-0,997006	-0,997042	-0,997012	-0,996977	-0,997714	-0,997714	-0,997714
12	5	PA7	32	Хорошая			-1,01593	-0,997065	-0,996266	-0,996976	-0,996976	-0,996976	-0,996976	-0,997786	-0,997782
13	6	PA7	32	Хорошая			-1,01667	-0,997042	-0,997042	-0,997042	-0,997788	-0,997799	-0,997714	-0,997751	-0,998521
14	329_1	PGC	PGC	Хорошая			-1	-1	-1	-1	-1	-1	-1	-1	-1
15	329_6	PGC	PGC	Хорошая			-1	-1	-1	-1	-1	-1	-1	-1	-1
16	331_1	PGC	PGC	Хорошая			-1	-1	-1	-1	-1	-1	-1	-1	-1
17	331_6	PGC	PGC	Хорошая			-1	-1	-1	-1	-1	-1	-1	-1	-1
18	s1:0,1	PGC	PGC	Плохая			-1	-1	-1	-1	-1	-1	-1	-1	-1

ФИГ. 14А

8. Выбрать столбцы, содержащие идентификаторы классификации, и нажать Вторичный идентификатор наблюдения.

Импорт SIMCA — [Матрица X]

Файл Главная страница Редактировать Просмотреть

Добавить данные | Первичный идентиф. переменной | Вторичный идентиф. переменной | Идентификатор первичного наблюдения | **Вторичный идентификатор наблюдения** | Идентификатор класса

Количественные | Качественные | Дата/Время | Переменная X | Переменная Y | Исключить | Включить

Типы данных | Роли переменных | Исключить

Матрица X

Первичн. идент.	Первичн. идент.	Вторичн. идент.	Вторичн. идент.	Вторичн. идент.	Вторичн. идент.	Вторичн. идент.	7	8	9	10	11	12
	Партия_цикл	Область обработки	Заполнение колески	Хорошая/Плохая	Бифуркация?	Сильное затягивание?	0	0,02	0,04	0,06	0,08	
2		1 PA7		32 Хорошая			-1,02032	-0,997782	-0,997051	-0,997075	-0,997782	-0,9978
3		2 PA7		32 Хорошая			-1,01667	-0,997811	-0,997026	-0,997714	-0,997714	-0,9977
4		3 PA7		32 Хорошая			-1,01593	-0,997783	-0,996976	-0,997783	-0,997783	-0,9977
5		4 PA7		32 Хорошая			-1,01742	-0,997045	-0,997041	-0,997034	-0,997781	-0,9977
6		5 PA7		32 Хорошая			-1,01593	-0,997758	-0,997791	-0,997791	-0,997791	-0,9977
7		6 PA7		32 Хорошая			-1,01525	-0,99785	-0,997111	-0,997111	-0,997783	-0,99
8		1 PA7		32 Хорошая			-1,02035	-0,997844	-0,997072	-0,997844	-0,997777	-0,9977
9		2 PA7		32 Хорошая			-1,01742	-0,997781	-0,997041	-0,997781	-0,997731	-0,9977
10		3 PA7		32 Хорошая			-1,01742	-0,997041	-0,997041	-0,997041	-0,997041	-0,9977
11		4 PA7		32 Хорошая			-1,01667	-0,996975	-0,997006	-0,997042	-0,997012	-0,9969
12		5 PA7		32 Хорошая			-1,01593	-0,997065	-0,996266	-0,996976	-0,996976	-0,9969
13		6 PA7		32 Хорошая			-1,01667	-0,997042	-0,997042	-0,997042	-0,997788	-0,9977
14	329_1	PCC	PCC	Хорошая			-1	-1	-1	-1	-1	
15	329_6	PCC	PCC	Хорошая			-1	-1	-1	-1	-1	
16	331_1	PCC	PCC	Хорошая			-1	-1	-1	-1	-1	
17	331_6	PCC	PCC	Хорошая			-1	-1	-1	-1	-1	
18	s1:0,1	PCC	PCC	Плохая			-1	-1	-1	-1	-1	
19	s1:0,12	PCC	PCC	Плохая			-1	-1	-1	-1	-1	
20	s2:0,5	PCC	PCC	Плохая			-1	-1	-1	-1	-1	
21	s2:0,4	PCC	PCC	Плохая			-1	-1	-1	-1	-1	
22	s2:0,35	PCC	PCC	Плохая			-1	-1	-1	-1	-1	
23	tm1:0,65	PCC	PCC	Плохая			-1	-1	-1	-1	-1	
24	tm1:0,68	PCC	PCC	Плохая			-1	-1	-1	-1	-1	
25	tm2:2	PCC	PCC	Плохая			-1	-1	-1	-1	-1	
26	tm2:1,75, s2:0,35	PCC	PCC	Плохая			-1	-1	-1	-1	-1	
27	tm1:0,78, s1:0,1	PCC	PCC	Плохая			-1	-1	-1	-1	-1	

4 III

Вопросы

9. Убедитесь, что все столбцы, содержащие значения объема, выбраны в качестве переменных X.

Импорт SIMCA [Матрица X]															
Идентификатор первичного наблюдения			Идентификаторы наблюдения			Типы данных			Роли переменных		Исключить		Автоформатирование		Импорт
Идентификатор первичного наблюдения	Вторичный идентификатор	Идентификатор класса	Количественные	Качественные	Дата/Время	Переменная X	Переменная Y	Исключить	Включить	Очистить форматиру	Применить форматиру	Завершить импорт			
42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	
0,7	0,72	0,74	0,76	0,78	0,8	0,82	0,84	0,86	0,88	0,9	0,92	0,94	0,96	0,98	
813134	-0,771643	-0,682579	-0,404407	0,240945	0,920306	1	1	1	1	1	1	1	1	1	1
804179	-0,7633	-0,682857	-0,432052	0,179919	0,941744	1	1	1	1	1	1	1	1	1	1
799459	-0,756732	-0,671411	-0,409283	0,219025	0,970041	1	1	1	1	1	1	1	1	1	1
777148	-0,726855	-0,610504	-0,270584	0,4299	1	1	1	1	1	1	1	1	1	1	1
765666	-0,715754	-0,597959	-0,251072	0,453664	1	1	1	1	1	1	1	1	1	1	1
768208	-0,719212	-0,624938	-0,320685	0,345817	1	1	1	1	1	1	1	1	1	1	1
804741	-0,755596	-0,615453	-0,216251	0,445688	1	1	1	1	1	1	1	1	1	1	1
0,80636	-0,763546	-0,652518	-0,318478	0,277313	1	1	1	1	1	1	1	1	1	1	1
792687	-0,744492	-0,616817	-0,239288	0,404626	1	1	1	1	1	1	1	1	1	1	1
776074	-0,71661	-0,578371	-0,169576	0,491405	1	1	1	1	1	1	1	1	1	1	1
774262	-0,722537	-0,584497	-0,202146	0,437634	1	1	1	1	1	1	1	1	1	1	1
772455	-0,712532	-0,565237	-0,205214	0,509624	1	1	1	1	1	1	1	1	1	1	1
977782	-0,912126	-0,729329	-0,350695	0,213061	0,764994	1	1	1	1	1	1	1	1	1	1
954412	-0,840881	-0,56747	-0,0843333	0,509679	0,938466	1	1	1	1	1	1	1	1	1	1
954412	-0,840881	-0,56747	-0,0843333	0,509679	0,938466	1	1	1	1	1	1	1	1	1	1
954412	-0,840881	-0,56747	-0,0843333	0,509679	0,938466	1	1	1	1	1	1	1	1	1	1
062192	0,303867	0,537772	0,742524	0,897116	0,984436	1	1	1	1	1	1	1	1	1	1
288777	0,485951	0,666355	0,817468	0,927978	0,989179	1	1	1	1	1	1	1	1	1	1
966603	-0,886968	-0,69617	-0,351367	0,0997823	0,480992	0,591903	0,591986	0,592261	0,593086	0,595306	0,600647	0,612118	0,634044	0,671169	0,711169
961686	-0,862025	-0,613039	-0,154799	0,437735	0,90469	1	1	1	1	1	1	1	1	1	1
961686	-0,862025	-0,613039	-0,154799	0,437735	0,90469	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
961686	-0,862025	-0,613039	-0,154799	0,437735	0,90469	1	1	1	1	1	1	1	1	1	1
961686	-0,862025	-0,613039	-0,154799	0,437735	0,90469	1	1	1	1	1	1	1	1	1	1
452298	0,67054	0,846233	0,960397	1	1	1	1	1	1	1	1	1	1	1	1

28/67

ФИГ. 15А

Импорт SIMCA — [Матрица X]

Файл Главная страница Редактировать Просмотреть

Добавить данные

Первичный идентиф. переменной
Идентификаторы переменных

Вторичный идентиф. переменной
Идентификаторы переменных

Идентификатор первичного наблюдения
Идентификаторы наблюдения

Вторичный идентификатор наблюдения
Идентификаторы наблюдения

Идентификатор класса

Количественные

Матрица X x

Первичн. идент.	Первичн. идент.	Вторичн. идент.	Вторичн. идент.	Вторичн. идент.	Вторичн. идент.	Вторичн. идент.
Первичн. идент.	Партия_цикл	Область обработки	Заполнение колонки	Хорошая/Плохая	Бифуркация?	Сильное затягивание?
2		1 PA7		32 Хорошая		
3		2 PA7		32 Хорошая		
4		3 PA7		32 Хорошая		
5		4 PA7		32 Хорошая		
6		5 PA7		32 Хорошая		
7		6 PA7		32 Хорошая		
8		1 PA7		32 Хорошая		
9		2 PA7		32 Хорошая		
10		3 PA7		32 Хорошая		
11		4 PA7		32 Хорошая		
12		5 PA7		32 Хорошая		
13		6 PA7		32 Хорошая		
14	329_1	PGC	PGC	Хорошая		
15	329_6	PGC	PGC	Хорошая		
16	331_1	PGC	PGC	Хорошая		
17	331_6	PGC	PGC	Хорошая		
18	s1:0,1	PGC	PGC	Плохая		
19	s1:0,12	PGC	PGC	Плохая		
20	s2:0,5	PGC	PGC	Плохая		
21	s2:0,4	PGC	PGC	Плохая		
22	s2:0,35	PGC	PGC	Плохая		
23	tm1:0,65	PGC	PGC	Плохая		
24	tm1:0,68	PGC	PGC	Плохая		
25	tm2:2	PGC	PGC	Плохая		
26	tm2:1,75, s2:0,35	PGC	PGC	Плохая		
27	tm1:0,78, s1:0,1	PGC	PGC	Плохая		

III

Вопрос

► Решение всех

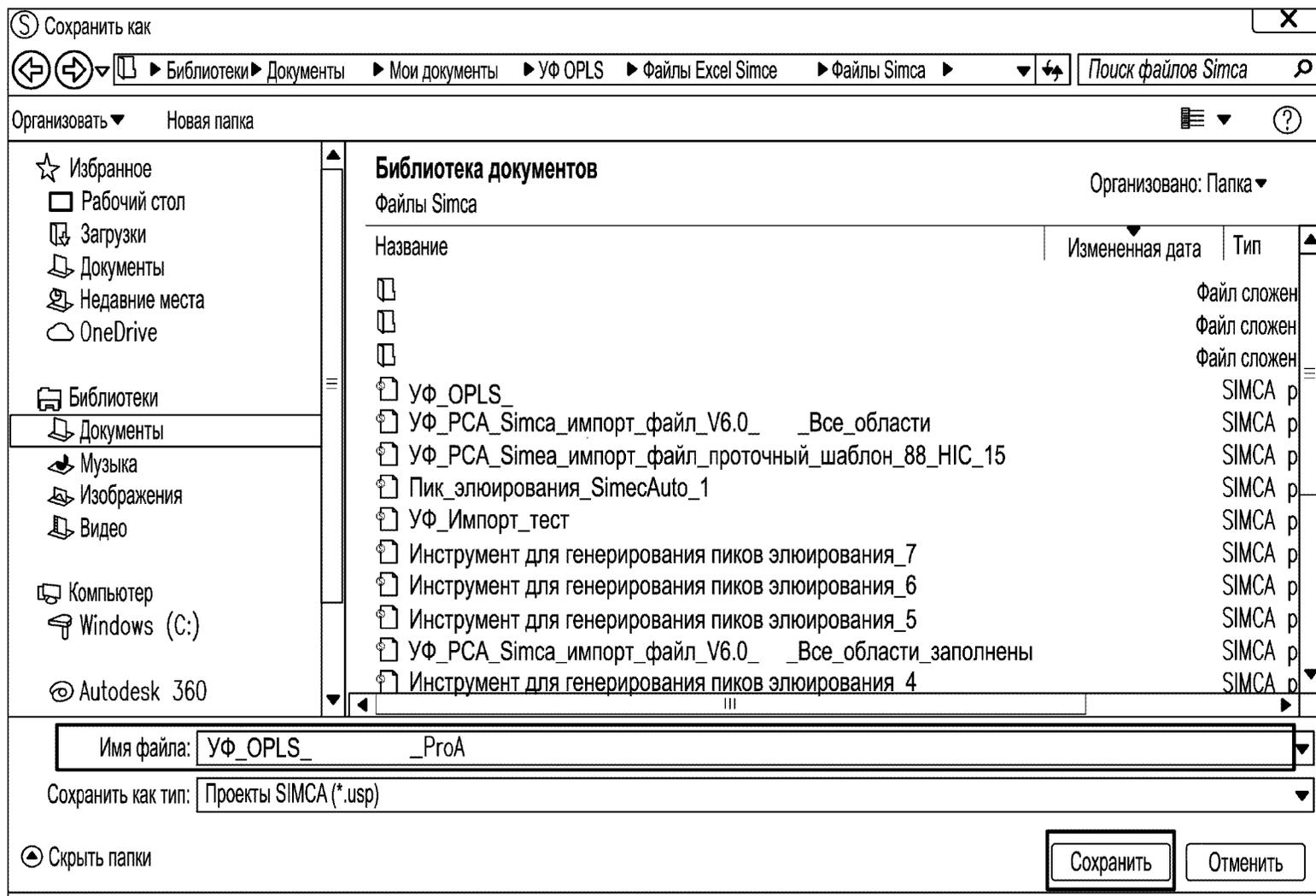
Лист: Матрица X (151 переменных, 26 наблюдений)

Вопросы отсутствуют

Продолж. На
ФИГ. 13В-2

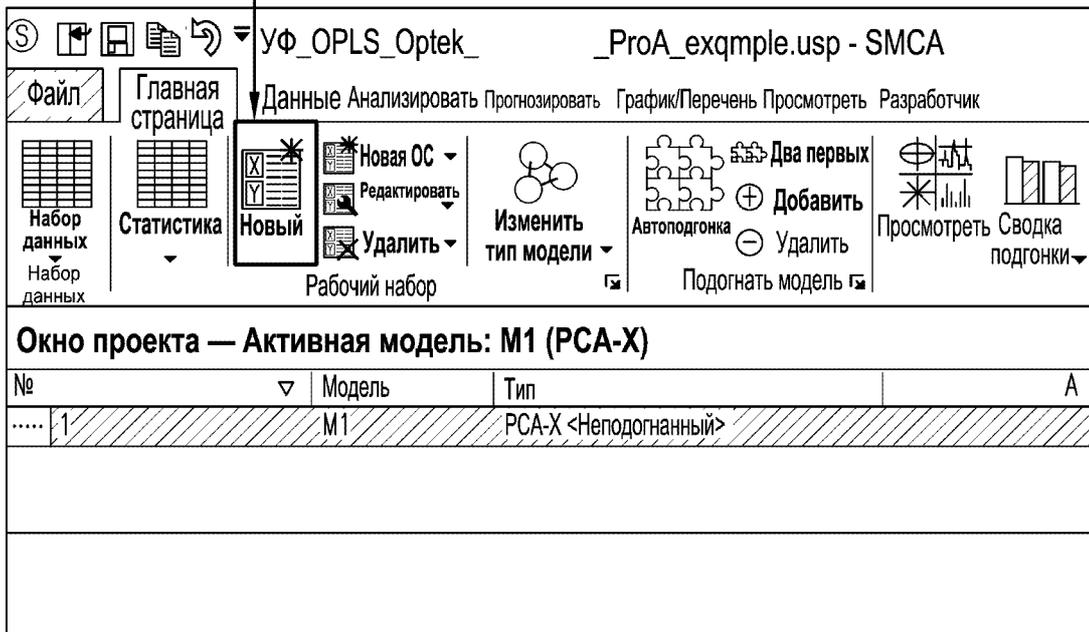
10. Проверьте нижнюю часть окна импорта на наличие раздела «Проблемы». Убедитесь, что он отображает «Вопросы отсутствуют». Затем нажмите «Завершить импорт».

11. Сохраните и назовите новый файл Simca. Назовите файл, выберите местоположение, где файл будет сохранен, затем нажмите «Сохранить».

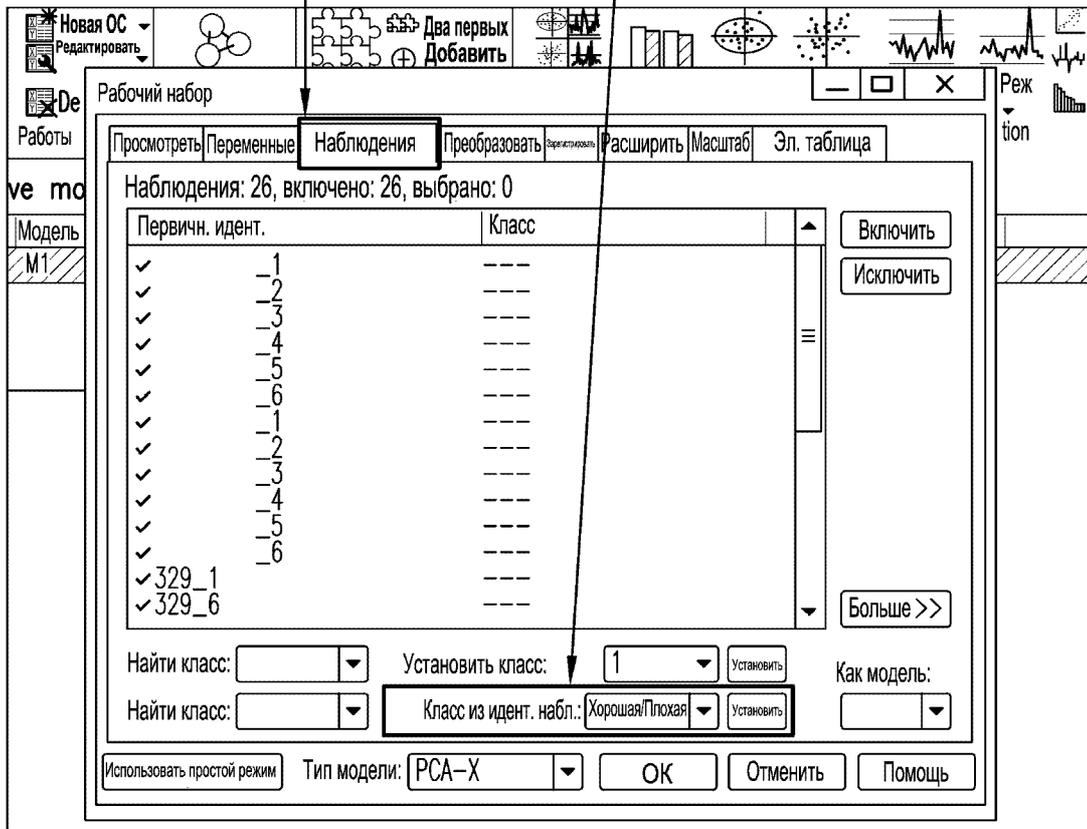


ФИГ. 16

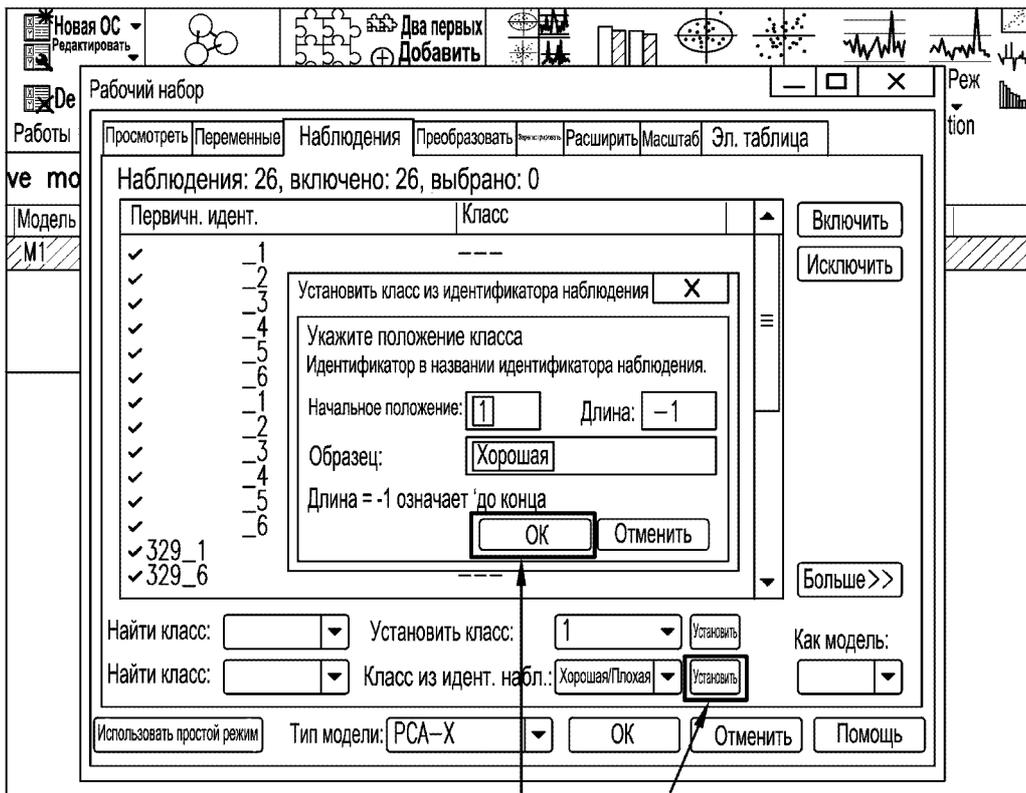
1 Выбрать новое



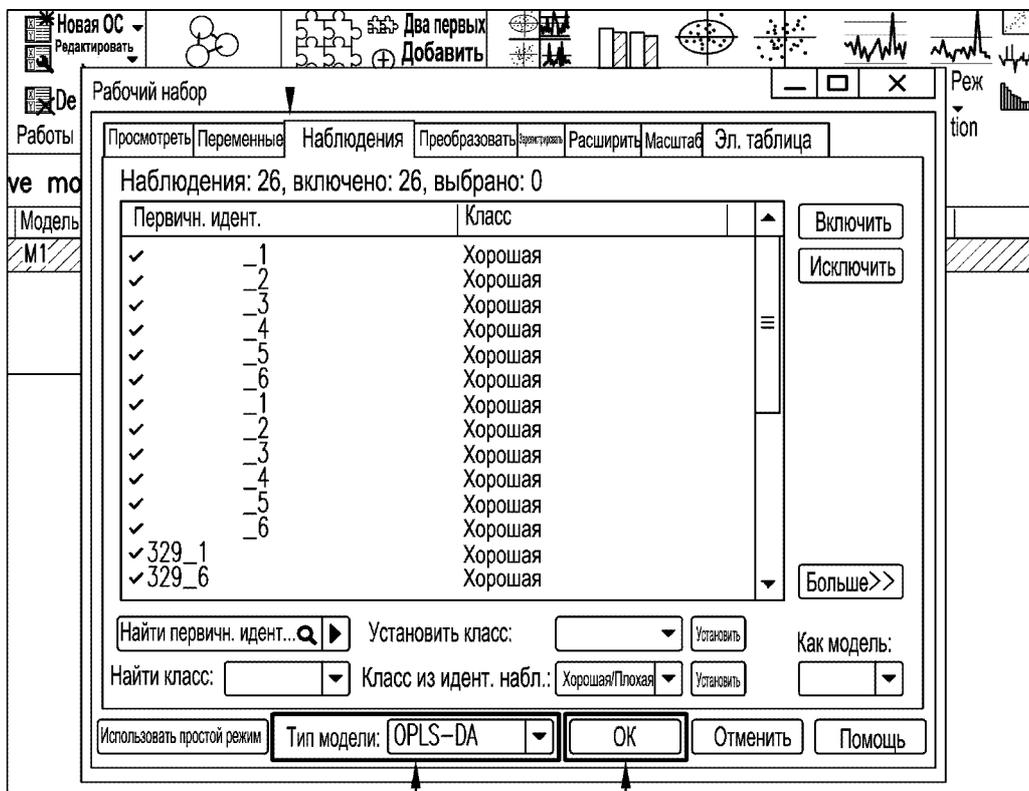
2. Выберите «Наблюдения», затем выберите стрелку вниз рядом с «Классом» из идент. набл. и выберите «Хорошая/Плохая».



ФИГ. 17



3. Выберите «установить», затем нажмите «OK» во всплывающем окне.



4. Тип модели отображается как OPLS-DA. Выбрать «OK» ФИГ. 18

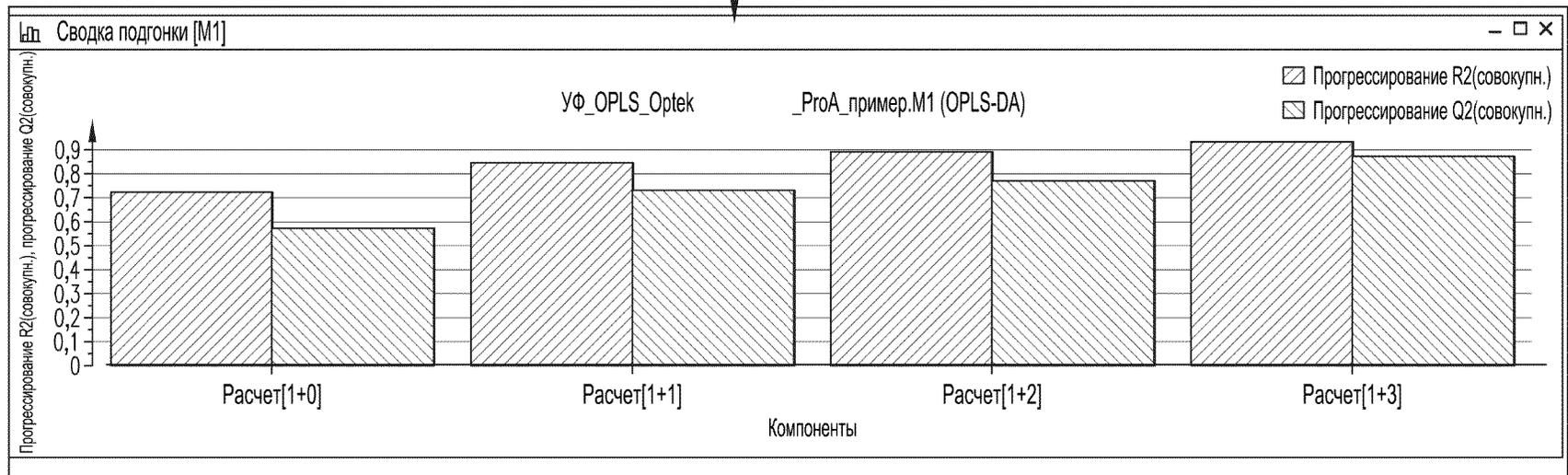
5. Нажмите Автоподгонку. Эта подгонка автоматически подгоняет оптимальное количество компонентов к модели. Внизу появится диаграмма, отображающая сводку подгонки

Файл Главная страница Данные Анализировать Прогнозировать Графики/Перечень Просмотреть Разработчик Инструменты

Набор данных | Статистика | Новый | Новая ОС | Редактировать | Удалить | Рабочий набор | Изменить тип модели | Автоподгонка | Два первых | Добавить | Удалить | Подогнать модель | Просмотреть | Сводка подгонки | Показатели | Загрузки | Гостиничный комплекс T2 | DMod | Наблюдаемое по сравн. с прогнозируемым | Коэффициенты | VIP

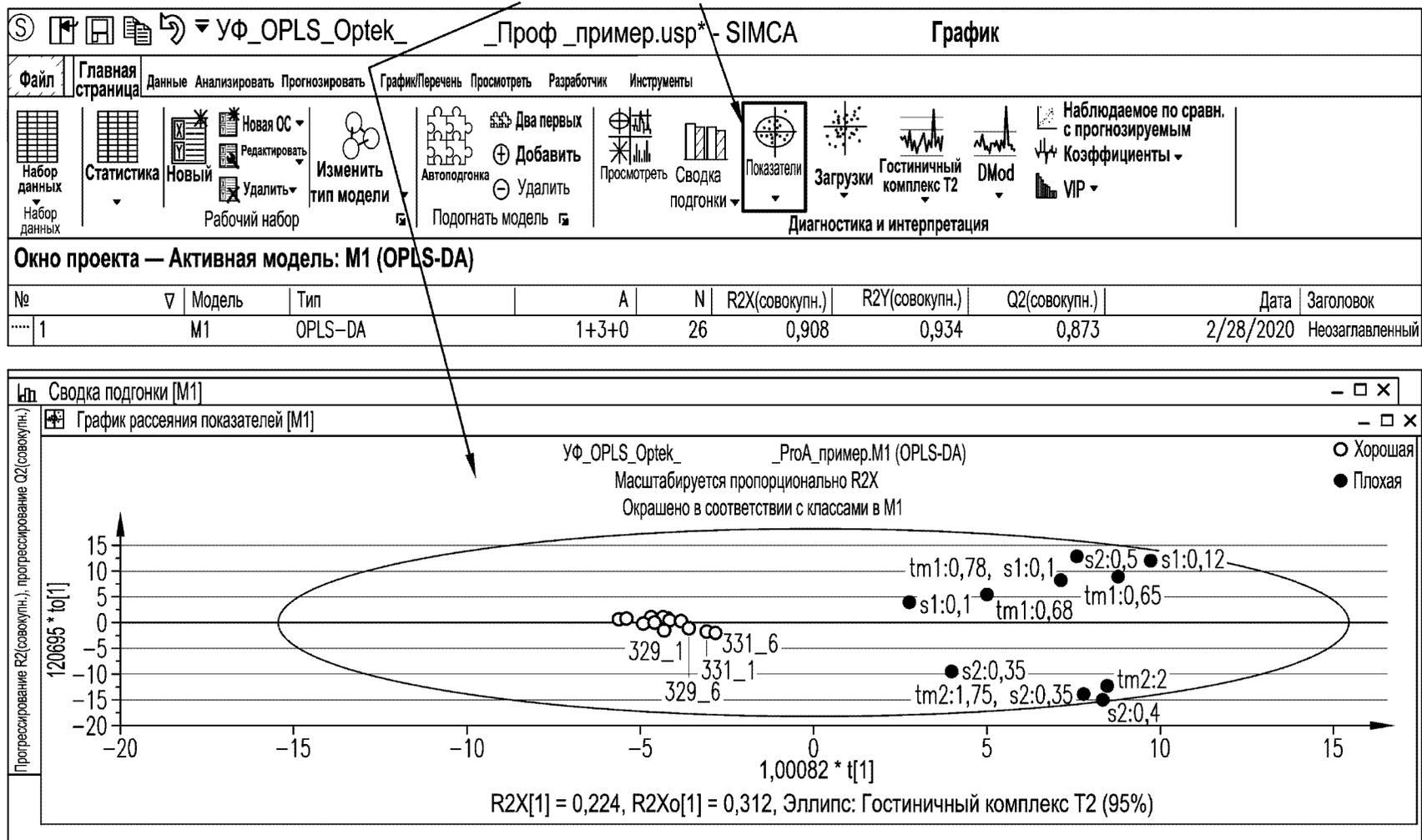
Окно проекта — Активная модель: M1 (OPLS-DA)

№	Модель	Тип	A	N	R2X(совокупн.)	R2Y(совокупн.)	Q2(совокупн.)	Дата	Заголовок
1	M1	OPLS-DA	1+3+0	26	0,908	0,934	0,873	2/28/2020	Незаглавленный



ФИГ. 19

6. Нажмите «Показатели». Это отобразит график рассеивания показателей, применяемый для интерпретации модели. Теперь модель OPLS сгенерирована.



ФИГ. 20

1. Перейдите в меню данных на верхней панели инструментов. Нажмите «Спектральные фильтры», затем «Производные».

УФ_OPLS_Optek ProA_пример.usp – SIMCA – [Рассеяние показателей... График

Файл Главная страница Данные Анализировать Прогнозировать Графики/Перечень Просмотреть Разработчик Инструменты

Импортировать набор данных Набор данных Объединить Разделить Транспонировать Удалить набор данных Генерировать переменные Локальное центрирование

Спектральные фильтры Фильтры временных рядов Сводка набора данных Карта отсутствующих значений Обзор обрезки Сводка

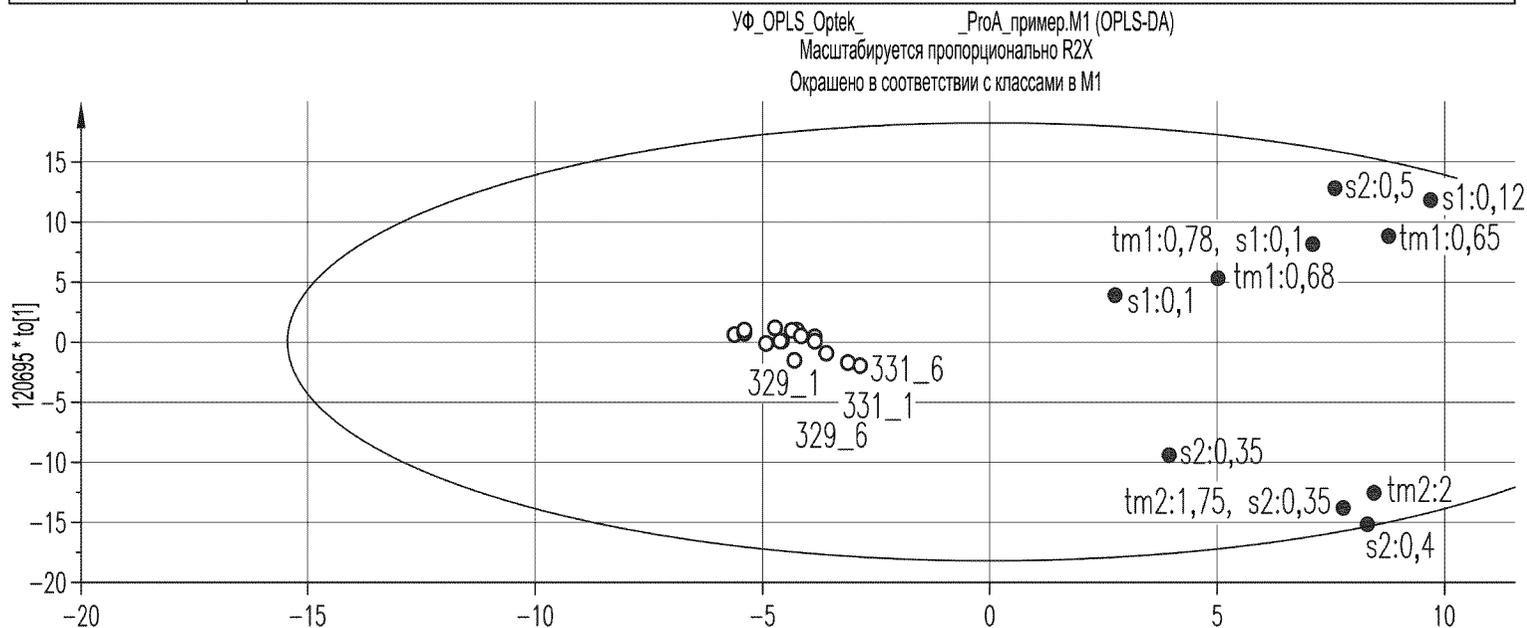
Иерархическая Неиерархическая Базовая модель

Окно проекта — Активная модель: M1 (OPLS-DA)

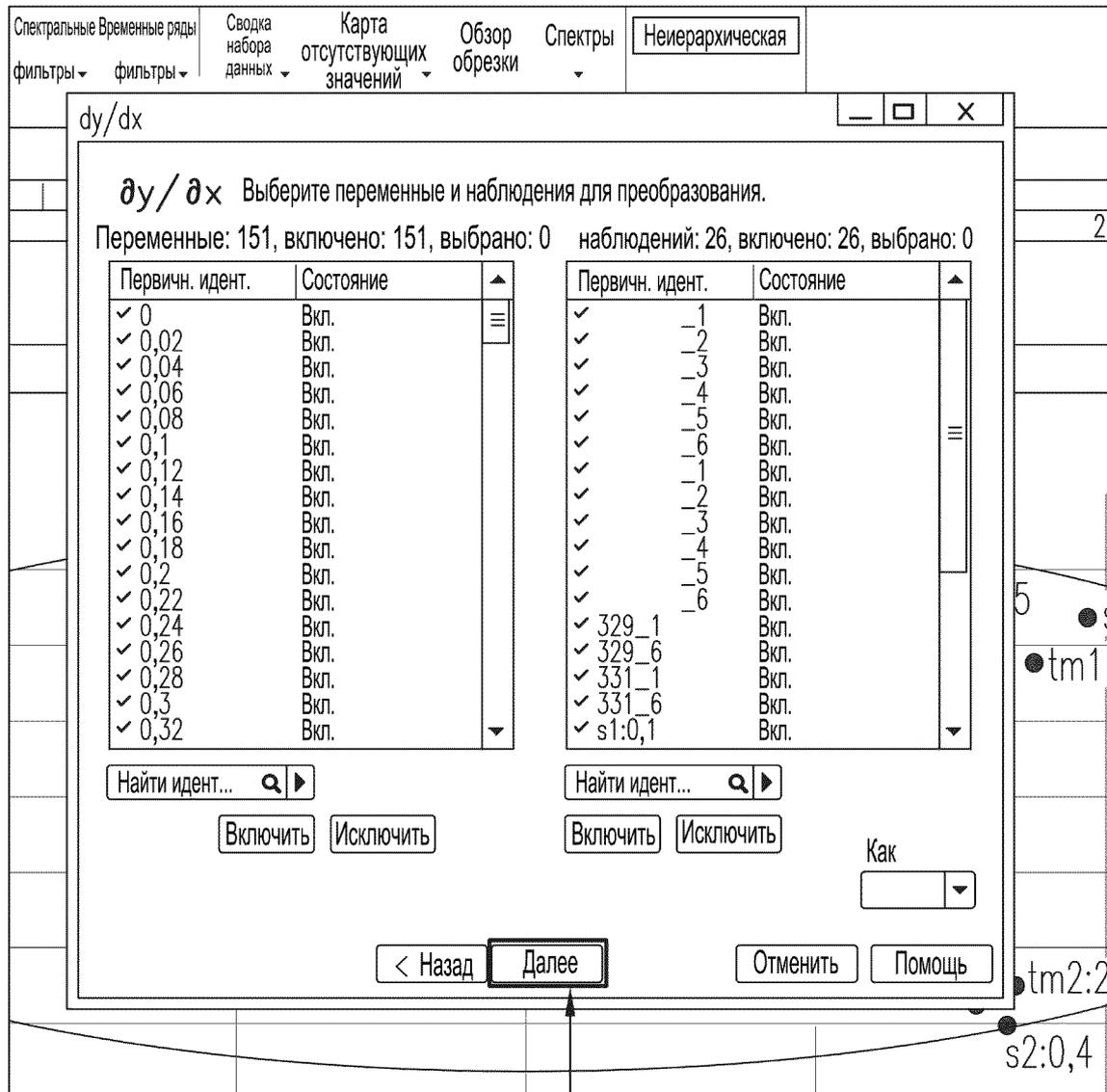
№	Модель	Тип	Q2(совокуп.)	Дата
1	M1	OPLS-DA	0,873	2/28/2020

Производные
Предварительно обработайте спектральные данные, используя первую, вторую или третью производные вместе с вариантом сглаживания.

График рассеяния показателей [M1] x



ФИГ. 21



2. Нажмите «Далее» во всплывающем меню.

ФИГ. 22А

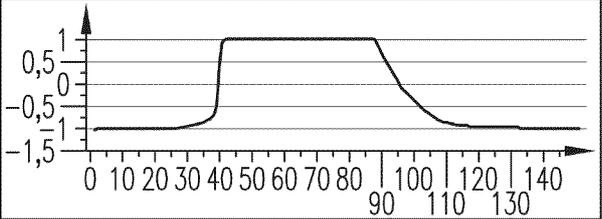
Спектральные Временные ряды
фильтры ▾ фильтры ▾

Сводка набора данных ▾ Карта отсутствующих значений ▾ Обзор обрезки ▾ Спектры ▾ Неиерархическая

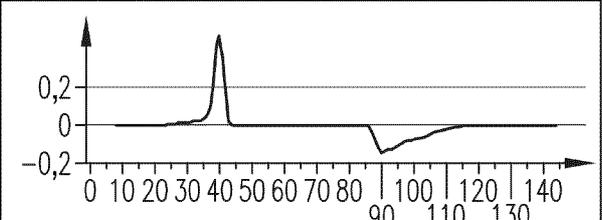
dy/dx

Задайте параметры для вывода. Выберите наблюдение в перечне, чтобы отобразить начальное и производное на графиках. Количество точек в каждой подмодели должно быть нечетным и составлять по меньшей мере 5

Начальные данные:



Производные данные:



Наблюдения:

Первичн. идент.
1
2
3
4
5
6
1
2
3
4
5
6
329_1
329_6
331_1
331_6

Производная спецификация

Производный порядок

- 1-я производная
- 2-я производная
- 3-я производная

Полиномиальный порядок

- Квадратичный
- Кубический

Точки в каждом

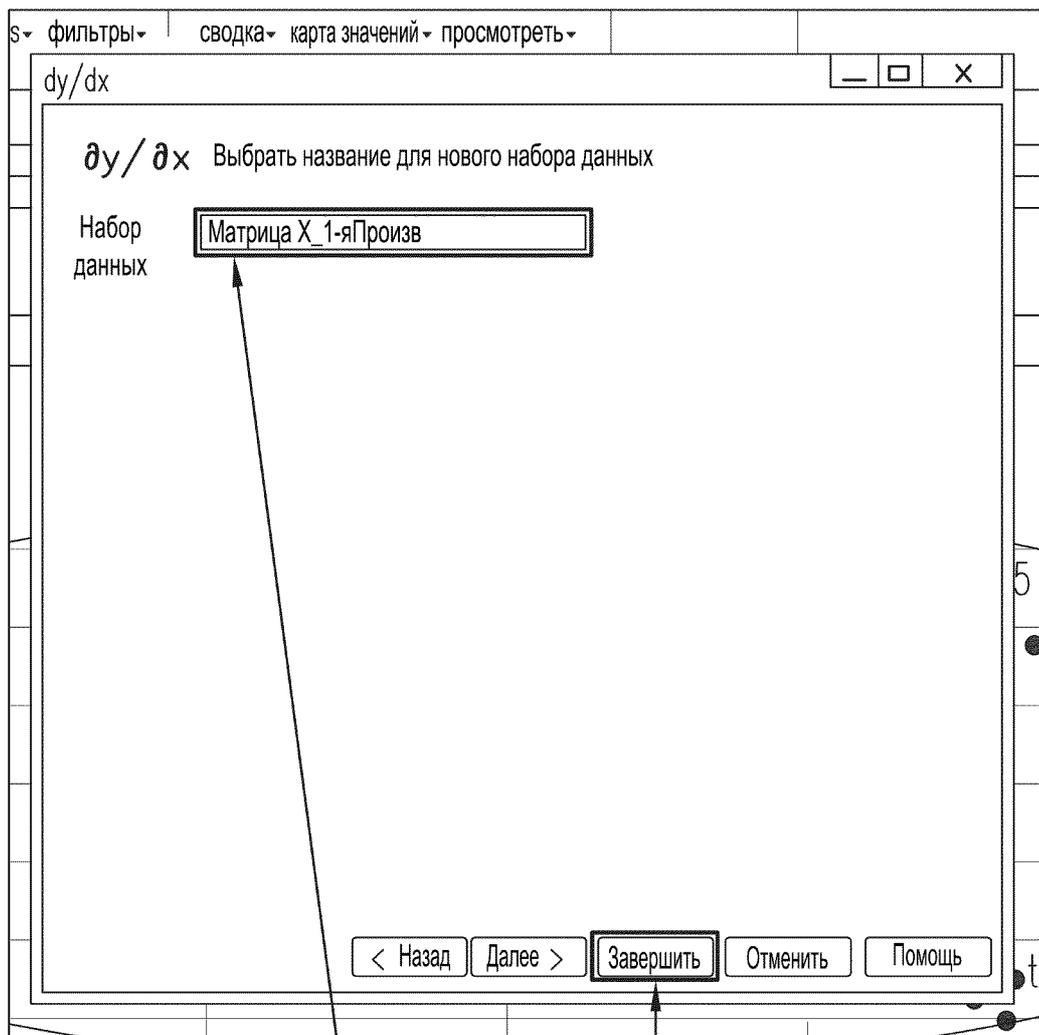
Расстояние между каждым

< Назад **Далее** > Отменить Помощь

tm2:2
s2:0,4

3. В следующем окне будут отображены данные для каждого прогона и то, как выглядит производная. Измените значение для Точек в каждом окне на 5, чтобы уменьшить уровень сглаживания. Затем нажмите «Далее»

ФИГ. 22В



4. Назовите новый набор данных, затем нажмите «Завершить».

ФИГ. 23А

5. Simca отобразит новый набор данных. Нажмите «Главная страница», затем «Новое», чтобы создать новую модель с набором данных.

The screenshot shows the SIMCA software interface. At the top, the 'Data' menu is open, with 'Главная страница' (Home page) and 'Новый' (New) highlighted. Below the menu, there are various icons for data management and analysis. The main window displays the project title 'Окно проекта — Активная модель: M1 (OPLS-DA)'. Below this is a table with columns for '№', 'Модель', 'Тип', 'A', 'N', and 'R2X(совокупн.)'. The table contains one row with the following values: 1, M1, OPLS-DA, 1+3+0, 26, 0,908. Below the table is a section titled 'График рассеяния показателей [M1]' and 'Набор данных- Матрица X_1-яПроизв'. This section contains a table with 11 columns and 19 rows of data.

	1	2	3	4	5	6	7	8	9	10	11
1	Первичн. идент.	Область обработки	Заполнение колонки	Хорошая/Плохая	Бифуркация?	Сильное затягивание?	\$1-яПроизв:0	\$1-яПроизв:0,0	\$1-яПроизв:0,0	\$1-яПроизв:0,0	\$1-яПроизв:0,0
2	\$Начальный идент.						0	0,02	0,04	0,06	0,08
3	1	PA7	32	Хорошая					00457751	2847e-05	00219482
4	2	PA7	32	Хорошая					00380105	4063e-05	00137627
5	3	PA7	32	Хорошая					00362867	6332e-05	00161266
6	4	PA7	32	Хорошая					00392843	00221062	00209224
7	5	PA7	32	Хорошая					00362539	5651e-06	00134397
8	6	PA7	32	Хорошая					00356697	1847e-05	00216681
9	1	PA7	32	Хорошая					00451382	0595e-05	00028699
10	2	PA7	32	Хорошая					00393717	4781e-05	00142676
11	3	PA7	32	Хорошая					00407518	00134492	00209832
12	4	PA7	32	Хорошая					00392512	9634e-07	00135213
13	5	PA7	32	Хорошая					00379879	2448e-05	00141966
14	6	PA7	32	Хорошая					00377659	00226009	00210148
15	329_1	PGC	PGC	Хорошая					0	0	0
16	329_6	PGC	PGC	Хорошая					0	0	0
17	331_1	PGC	PGC	Хорошая					0	0	0
18	331_6	PGC	PGC	Хорошая					0	0	0
19	s1:0,1	PGC	PGC	Плохая					0	0	0

40/67

ФИГ. 23В

6. Выбрать производный набор данных и отменить выбор начального набора данных

Указать, какие наборы данных включать и в каком порядке их включать

Набор данных	Размер (Вар./Набл.)
<input type="checkbox"/> Матрица X	151 x 26
<input checked="" type="checkbox"/> X матрица_1-яПроизв	151 x 26

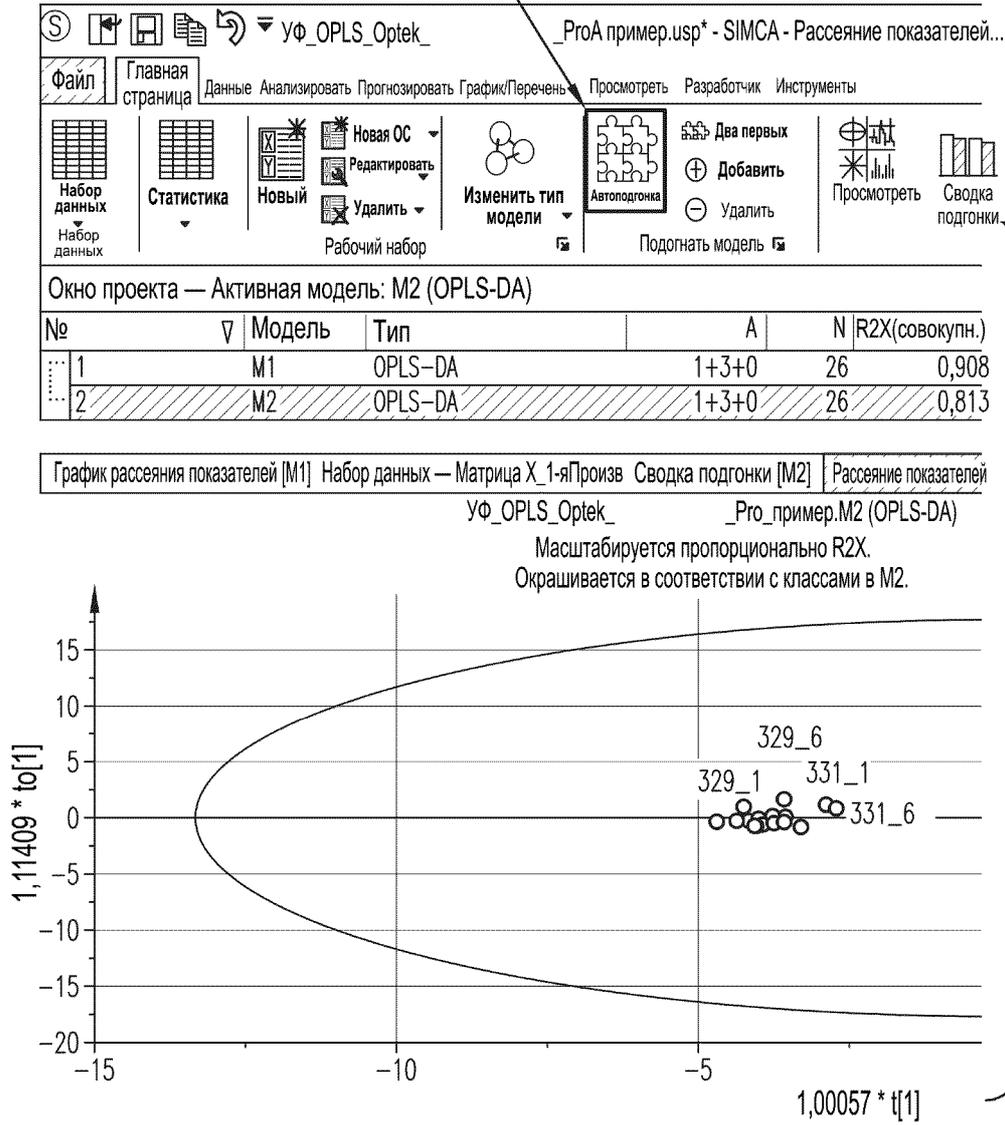
Сохранить как наборы данных по умолчанию Как модель:

Использовать простой режим Тип модели:

	10	11
\$1-яПроизв:0,0	\$1-яПроизв:0,0	
0,06	0,08	
2847e-05	00219482	
4063e-05	00137627	
6332e-05	00161266	
00221062	00209224	
5651e-06	00134397	
1847e-05	00216681	
0595e-05	00028699	
4781e-05	00142676	
00134492	00209832	
9634e-07	00135213	
00379879	2448e-05	00141966
00377659	00226009	00210148
PGC	0	0

ФИГ. 24А

8. Нажмите Автоподгонку, затем показатели. Теперь будет отображена модель OPLS с первой производной.



Продолж. На
ФИГ. 25-2

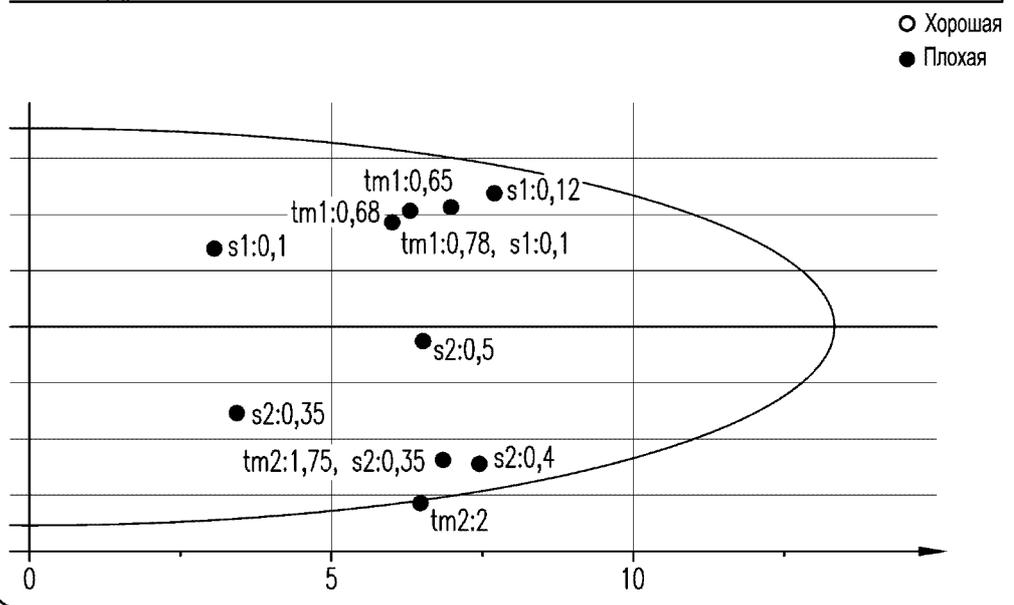
ФИГ. 25-1

8. Нажмите Автоподгонку, затем показатели. Теперь будет отображена модель OPLS с первой производной.

График

R2Y(совокупн.)	Q2(совокупн.)	Дата	Заголовок	Иерархическая
0,934	0,873	2/28/2020	Неозаглавленный	
0,958	0,894	3/3/2020	Неозаглавленный	

Продолж. От Фиг.25-1



ФИГ. 25-2

1. Перейдите во вкладку «Данные», затем нажмите «Импортировать набор данных»

The screenshot shows the SIMCA software interface. At the top, the title bar reads "УФ_OPLS_Optek_ _ProA пример.usp* - SIMCA - Рассеяние пок [Predicted Sc...График". The main menu bar includes "Файл", "Главная страница", and "Данные". The "Данные" menu is open, showing options: "Импортировать набор данных", "Набор данных", "Объединить", "Разделить", "Транспонировать", "Удалить набор данных", "Генерировать переменные", "Локальное центрирование", "Спектральные фильтры", "Фильтры временных рядов", and "sum". Below the menu, there are labels: "Набор данных", "Модифицировать набор данных", and "Фильтры".

The main window title is "Окно проекта — Активная модель: M2 (OPLS-DA)". Below this is a table with the following data:

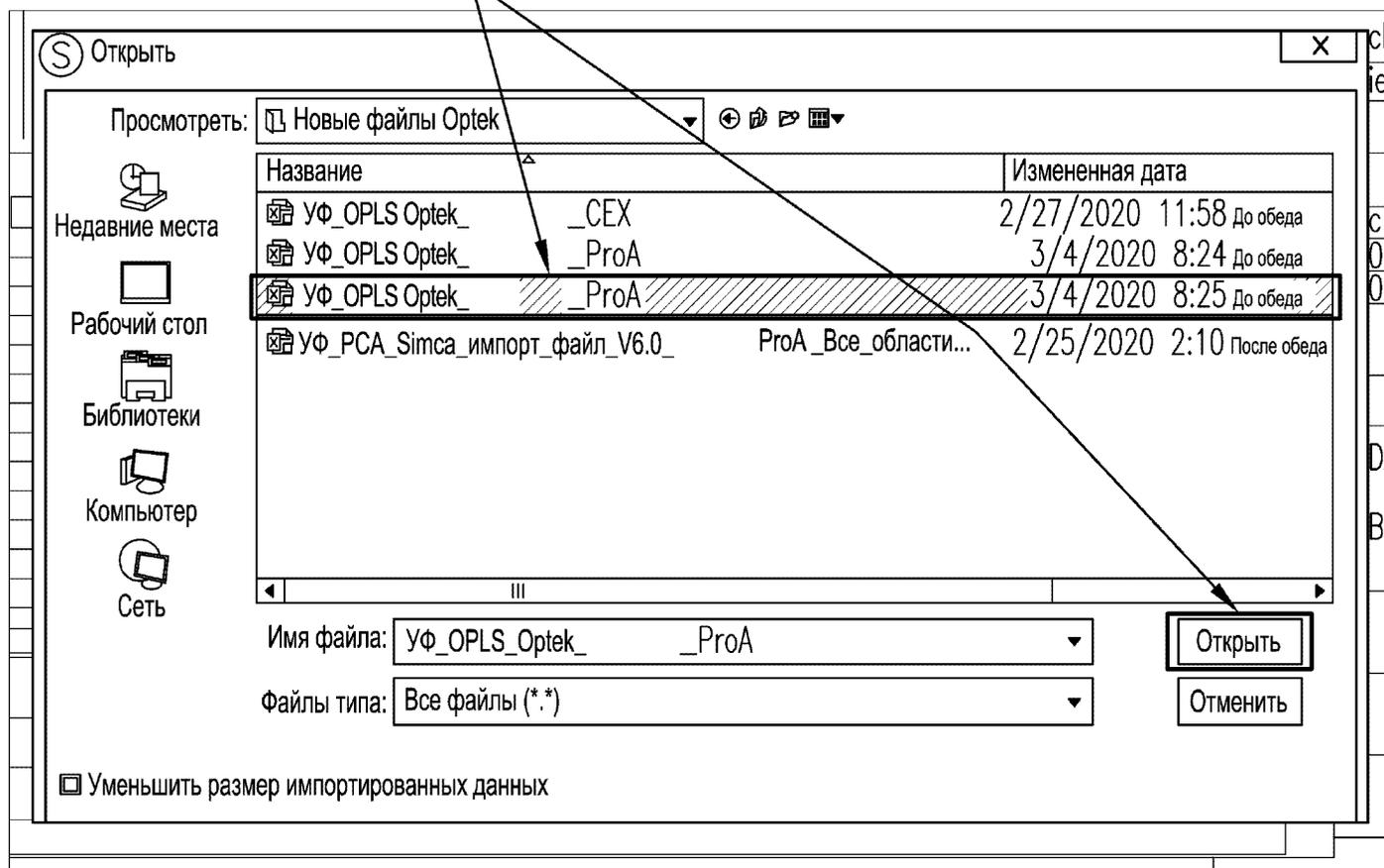
№	Модель	Тип	A
1	M1	OPLS-DA	1+3+0
2	M2	OPLS-DA	1+3+0

Below the table, there are several buttons: "Перестановка [M2]", "Сводка подгонки [M2]", "Набор данных — Матрица X 1-яПроизв", and "Рассеяние показателей".

At the bottom, there is a graph titled "UV_OPLS_Opt". The y-axis has labels 10 and 15. The graph shows a smooth, upward-sloping curve.

ФИГ. 26А

2. Выберите файл с новыми данными, которые будут прогнозированы моделью, и нажмите «Открыть».



ФИГ. 26В



3. Выберите лист, содержащий данные, и нажмите «OK».

ФИГ. 27А

4. Придерживайтесь того же способа, который описан ранее, чтобы отформатировать импорт. Переименуйте набор данных, чтобы он был уникальным (двойное нажатие), и исключите данные, уже находящиеся в другом наборе данных, выбрав строки данных и нажав «Исключить». Нажмите «Завершить» и импортируйте, когда закончите.

Импорт SIMCA — [Матрица X_335]

Файл Главная страница Редактировать Просмотреть

Добавить данные Первичный идентиф. переменной Вторичный идентиф. переменной Идентификаторы переменных

Идентификатор первичного наблюдения Вторичный идентификатор наблюдения Идентификаторы наблюдения

Идентификатор класса

Количественные Качественные Дата/Время

Переменная X Переменная Y

Исключить Включить

Очистить форматирование Применить форматирование Автоформатирование

Завершить импорт Импорт

Матрица X_335 x|⊕

	Первич. идент.	Вторичн. идент.	Вторичн. идент.	Вторичн. идент.	Вторичн. идент.	Вторичн. идент.	7	8	9	10	11	12	13	14	15
Первичн. идент.	Партия_цикл	Область обработки	Заполнение колонки	Хорошая/Плохая	Бифуркация?	Сильное затягивание?	0	0,02	0,04	0,06	0,08	0,1	0,12	0,14	0,16
2	8031100329_1	PA7	32	Хорошая			-1,02032	-0,997782	-0,997051	-0,997075	-0,997782	-0,997808	-0,997782	-0,998423	-0,998523
3	8031100329_2	PA7	32	Хорошая			-1,01667	-0,997811	-0,997026	-0,997714	-0,997714	-0,997714	-0,997714	-0,99782	-0,998521
4	8031100329_3	PA7	32	Хорошая			-1,01593	-0,997783	-0,996976	-0,997783	-0,997783	-0,997783	-0,997783	-0,998522	-0,998522
5	8031100329_4	PA7	32	Хорошая			-1,01742	-0,997045	-0,997041	-0,997034	-0,997781	-0,997781	-0,997714	-0,997781	-0,998453
6	8031100329_5	PA7	32	Хорошая			-1,01593	-0,997758	-0,997791	-0,997791	-0,997791	-0,997791	-0,998463	-0,998463	-0,998522
7	8031100329_6	PA7	32	Хорошая			-1,01525	-0,99785	-0,997111	-0,997111	-0,997783	-0,9978	-0,99785	-0,99785	-0,998589
8	8031100331_1	PA7	32	Хорошая			-1,02035	-0,997844	-0,997072	-0,997844	-0,997777	-0,997777	-0,99854	-0,998518	-0,998535
9	8031100331_2	PA7	32	Хорошая			-1,01742	-0,997781	-0,997041	-0,997781	-0,997731	-0,997728	-0,997781	-0,998453	-0,998453
10	8031100331_3	PA7	32	Хорошая			-1,01742	-0,997041	-0,997041	-0,997041	-0,997041	-0,997714	-0,997754	-0,997714	-0,997749
11	8031100331_4	PA7	32	Хорошая			-1,01667	-0,996975	-0,997006	-0,997042	-0,997012	-0,996977	-0,997714	-0,997714	-0,997714
12	8031100331_5	PA7	32	Хорошая			-1,01593	-0,997065	-0,996266	-0,996976	-0,996976	-0,996976	-0,996976	-0,997786	-0,997782
13	8031100331_6	PA7	32	Хорошая			-1,01667	-0,997042	-0,997042	-0,997042	-0,997788	-0,997799	-0,997714	-0,997751	-0,998521
14	8031100335_1	PA7	32				1,02172	0,997025	0,996368	0,997086	0,997086	0,997086	0,997086	0,997797	0,997761
15	8031100335_2	PA7	32				-1,0181	-0,99781	-0,996272	-0,99704	-0,996973	-0,99704	-0,997713	-0,997775	-0,997713
16	8031100335_3	PA7	32				-1,01742	-0,997041	-0,996301	-0,996352	-0,997022	-0,997041	-0,997041	-0,997781	-0,997781
17	8031100335_4	PA7	32				-1,01739	-0,997303	-0,997015	-0,997015	-0,997015	-0,997755	-0,997755	-0,997755	-0,997755
18	8031100335_5	PA7	32				-1,0181	-0,996973	-0,9963	-0,9963	-0,996973	-0,996993	-0,996973	-0,997731	-0,997713
19	8031100335_6	PA7	32				-1,01593	-0,997168	-0,997023	-0,997063	-0,996976	-0,997019	-0,997782	-0,99785	-0,997836

III

Вопросы

▶ Решение всех

Лист: Матрица X_335 (151 переменных, 6 наблюдений)

Вопросы отсутствуют

ФИГ. 27В

5. Перейдите во вкладку «Прогнозировать». Нажмите «Как набор данных», затем выберите импортированный набор данных.

The screenshot shows the SIMCA software interface. The top menu bar includes 'Главная страница', 'Данные', 'Анализировать', 'Прогнозировать', 'График/перечень', 'Просмотреть', 'Разработчик', and 'Инструменты'. The 'Прогнозировать' tab is active. A dropdown menu is open under 'Как набор данных', showing options: 'Набор данных', 'Матрица X', 'Матрица X_335', 'Набор прогнозирования', and 'Матрица X PS'. The 'Матрица X_335' option is selected. Below the menu, a table displays the data matrix structure:

№	Тип	A	N	R2X(совокупн.)	R2Y(совокупн.)	Q2(совокупн.)
1	OPLS-DA	1+3+0	26	0,908	0,934	0,873
2	OPLS-DA	1+3+0	26	0,813	0,958	0,894

Below the table, there are tabs for 'Набор прогнозирования - Матрица X_335 [M...', 'График прогнозируемого рассеяния показателей [M...', 'График прогнозируемого рассеяния колонок [...', and 'Набор данных - Матрица X_335 *'. A data matrix table is displayed below the tabs:

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
1	Первичн. идент.	Область обработки	Заполнение колонки	Хорошая/Плохая	Бифуркация?	Сильное затягивание?	0	0,02	0,04	0,06	0,08	0,1	0,12	0,14	0,16	0,18
2	1	PA7	32				-1,02172	0,997025	0,996368	0,997086	0,997086	0,997086	0,997086	0,997797	0,997761	-0,9
3	2	PA7	32				-1,0181	-0,99781	0,996272	-0,99704	0,996973	-0,99704	0,997713	0,997775	0,997713	0,99
4	3	PA7	32				-1,01742	0,997041	0,996301	0,996352	0,997022	0,997041	0,997041	0,997781	0,997781	0,99
5	4	PA7	32				-1,01739	0,997303	0,997015	0,997015	0,997015	0,997755	0,997755	0,997755	0,997755	0,99
6	5	PA7	32				-1,0181	0,996973	-0,9963	-0,9963	0,996973	0,996993	0,996973	0,997731	0,997713	0,99
7	6	PA7	32				-1,01593	0,997168	0,997023	0,997063	0,996976	0,997019	0,997782	-0,99785	0,997836	0,99

49/67

ФИГ. 28

6. Нажмите «Показатели PS». График показателей, показывающий прогнозируемые показатели, появится в окне внизу.

УФ_OPLS_Optek_ _ProA пример.usp* - SIMCA – [прогнозируемое рас...]

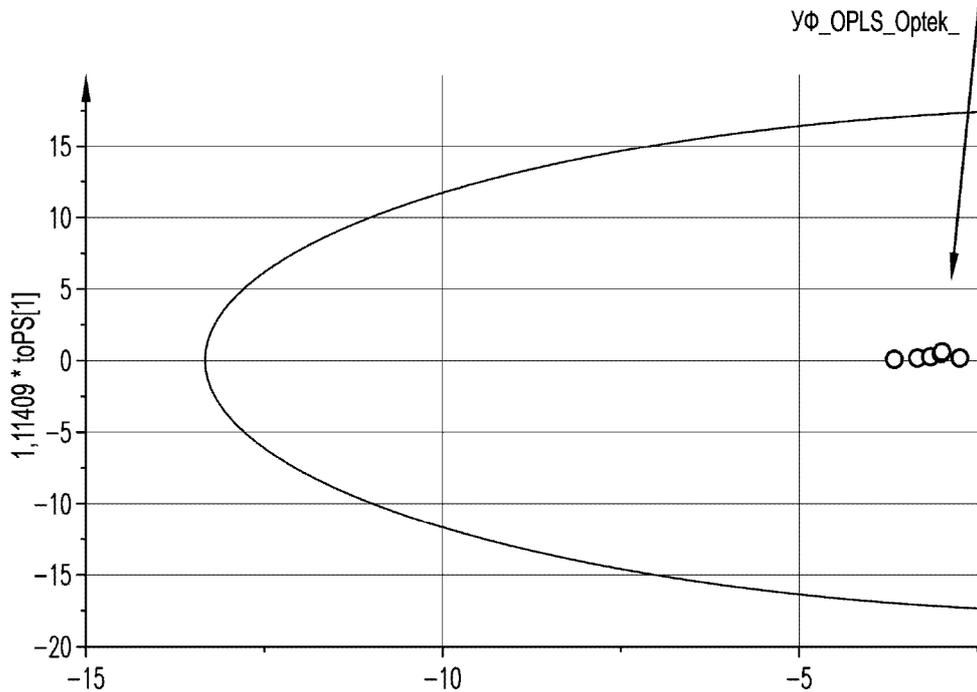
Файл Главная страница Данные Анализировать Прогнозировать График/перечень Просмотреть Разработчик Инструменты

Указать Как набор данных Как рабочий набор Рабочий набор дополнений Класс Удалить набор прогнозирования Удалить набор прогнозирования Перечень прогнозирования Перечень Y PS Показатели PS Гостиничный комплекс T2PS DMod PS Графики

Окно проекта — Активная модель: M2 (OPLS-DA)

№	Модель	Тип	A	N
1	M1	OPLS-DA	1+3+0	26
2	M2	OPLS-DA	1+3+0	26

Набор прогнозирования - Матрица X PS [M2] График прогнозируемого рассеяния показателей [M... График прогнозируемого рассеяния колонок [...]



Продолж. На
ФИГ. 29-2

ФИГ. 29-1

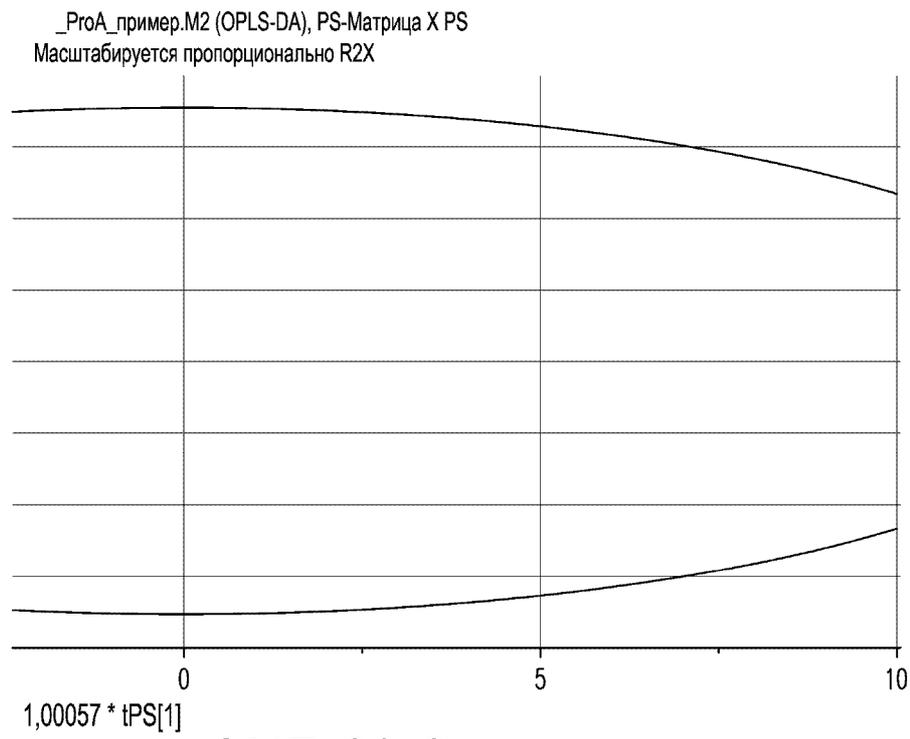
Продолж. От
Фиг. 29-1

График

 Контрольные диаграммы PS
 PS вклада
 Временной ряд PS
 ROC
 График Кумана
 Перечень классификации
 Таблица ошибочной классификации
 Классификация
 Что-если
 Инструменты

R2X(совокупн.)	R2Y(совокупн.)	Q2(совокупн.)	Дата	Заголовок
0,908	0,934	0,873	2/28/2020	Незаглавленный
0,813	0,958	0,894	3/3/2020	Незаглавленный

Набор данных — Матрица X 335 График прогнозируемого рассеяния показателей [M... График прогнозируемого рассеяния показателей [M...



ФИГ. 29-2

7. Чтобы добавить точки из модели в график показателей для сравнения того, где находятся прогнозируемые точки, нажмите «Указать», измените источник на начальный набор данных, выберите наблюдения данных, с которыми хотите сравнить, нажмите стрелку вправо и нажмите «ОК».

The screenshot shows a software interface with a menu bar, a toolbar, and a main window. The main window contains a plot of $1,11409 * \text{topS}[1]$ versus an unlabeled x-axis. A dialog box titled "Указать Набор Прогнозирования" is open, showing a list of available observations and a list of selected observations. The dialog box also includes a search field, a "Найти..." button, and buttons for "Удалить" and "Удалить все".

The main window displays the following data:

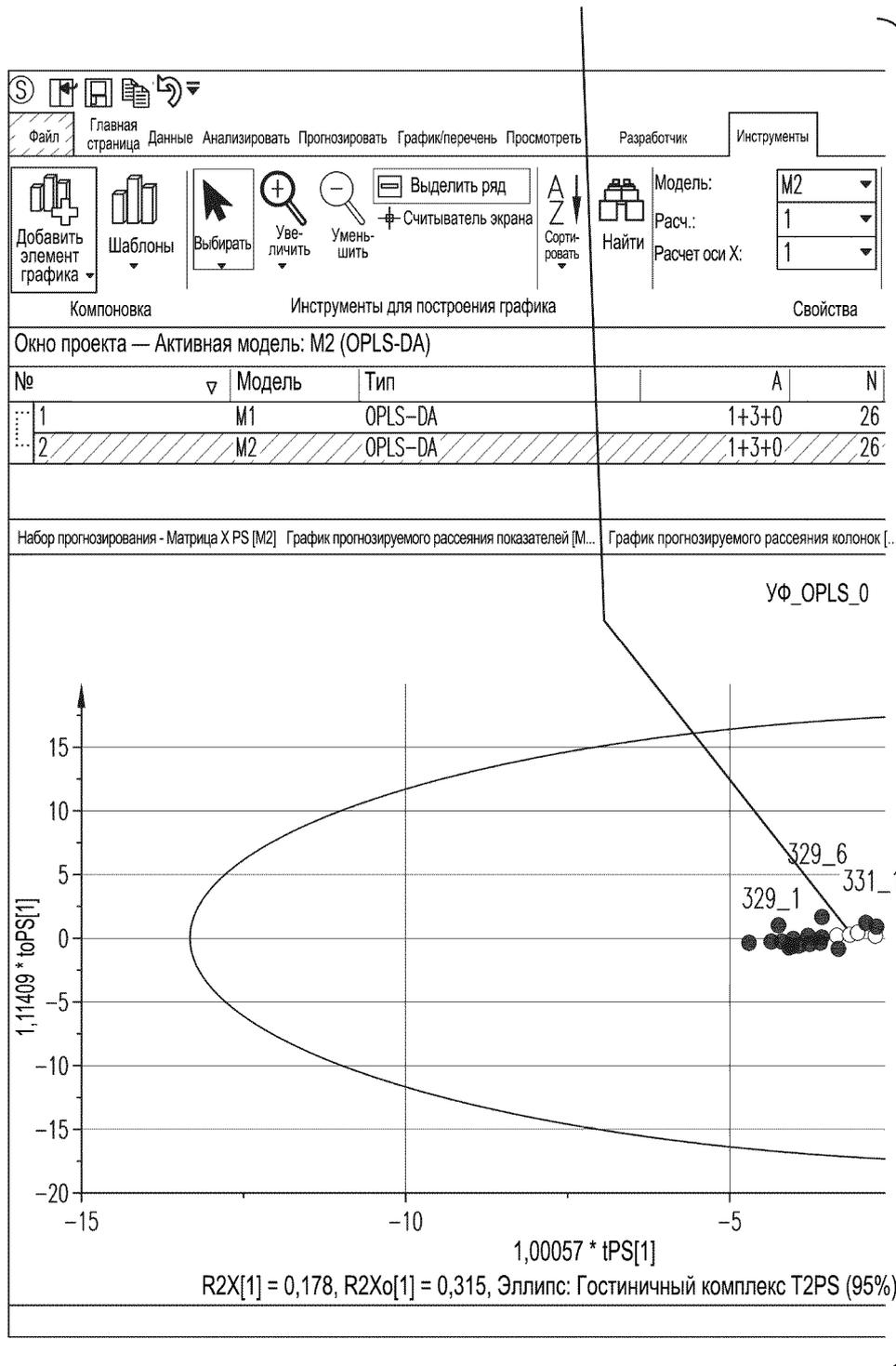
№	Модель	Тип
1	M1	OPLS-DA
2	M2	OPLS-DA

The plot shows a curve representing the relationship between the variables. The y-axis ranges from -15 to 15, and the x-axis ranges from 0 to 15. The curve starts at approximately (0, 0), rises to a peak of about 15 at x ≈ 10, and then falls back to about -15 at x ≈ 15.

52/67

FIG.30

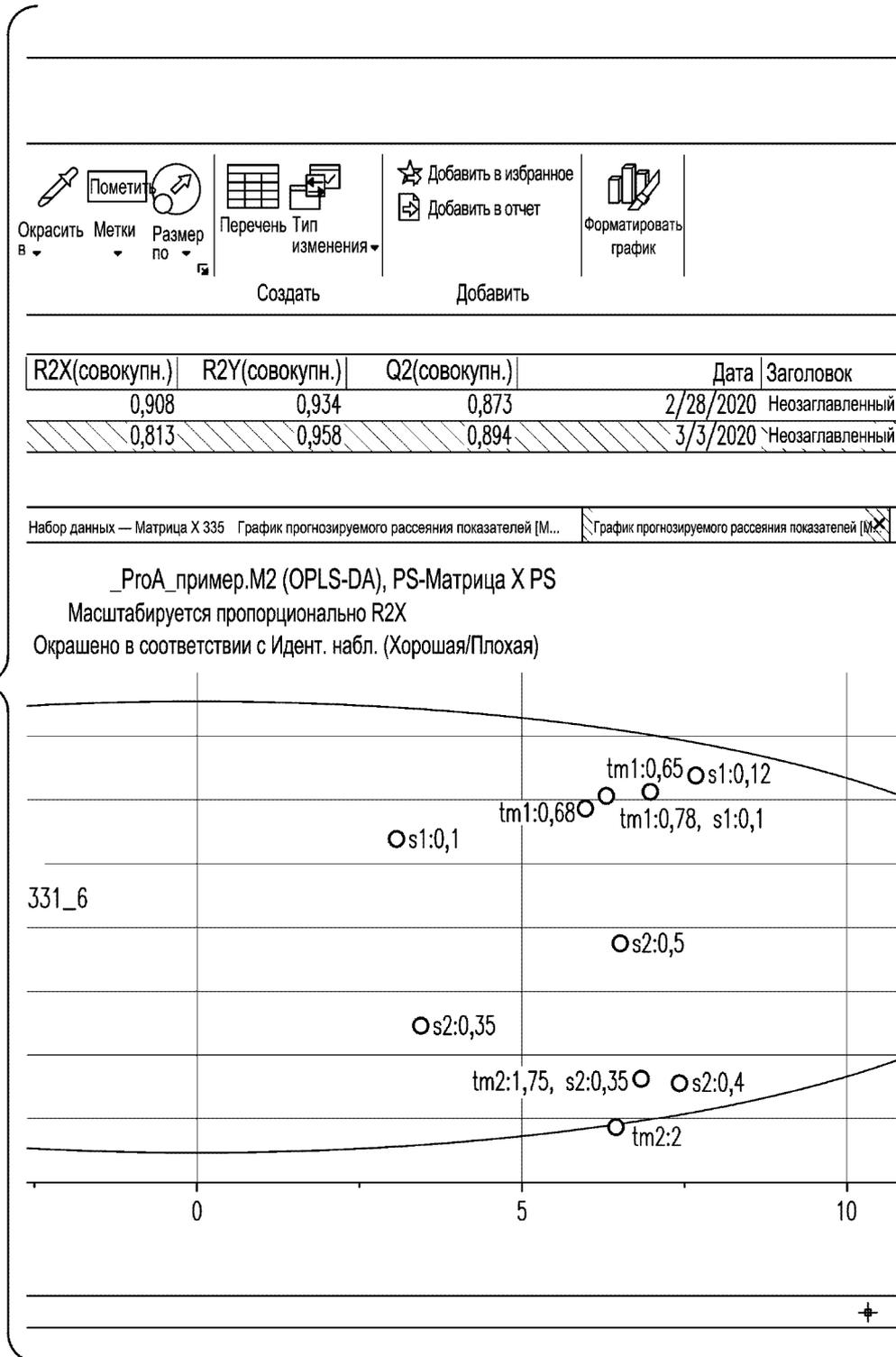
8. На графике прогнозируемых показателей теперь будут отображаться точки из графика обычных показателей. Данные окрашены в зависимости от категоризации «Хорошая/Плохая». Поскольку новые прогнозируемые прогоны не были классифицированы как хорошие или плохие, они окрашены в серый.



Продолж, на
ФИГ. 31-2

ФИГ. 31-1

Продолж. От
Фиг. 31-1

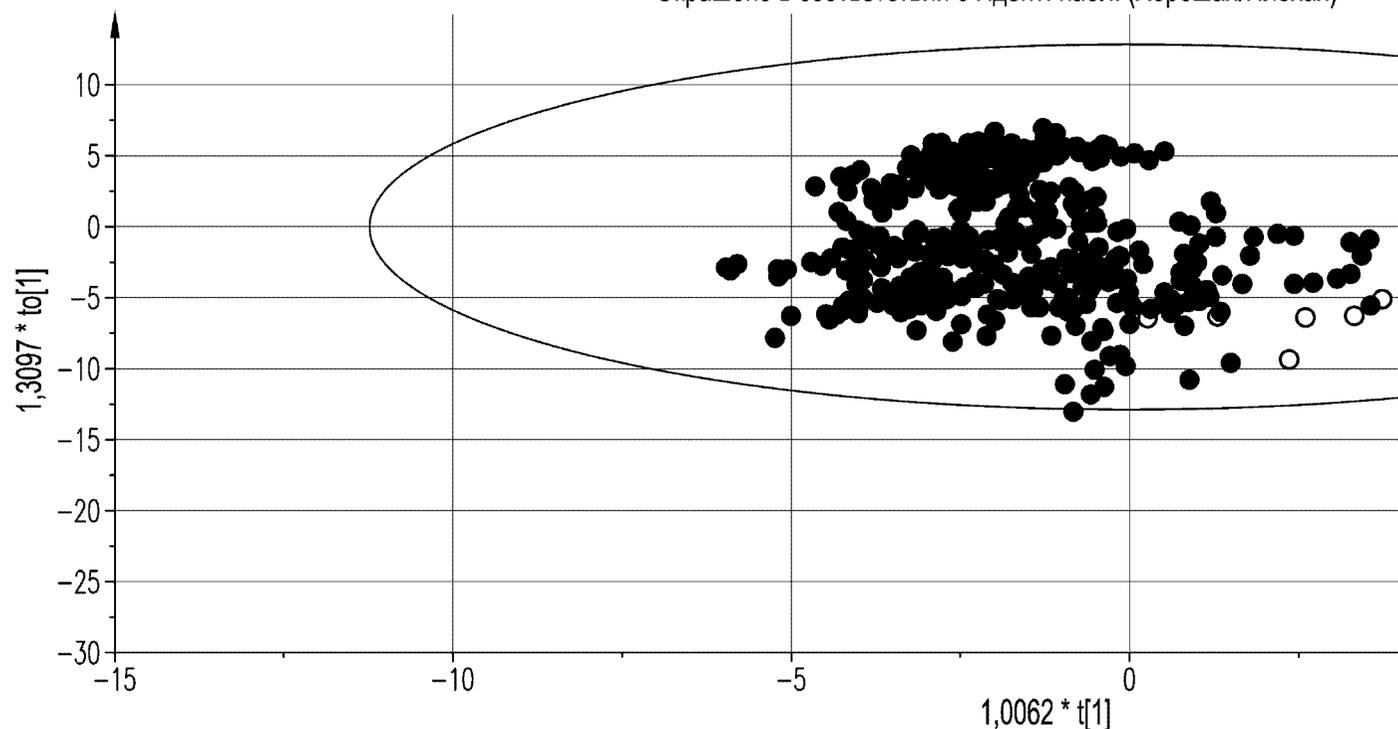


ФИГ. 31-2

№	▼ Модель	Тип	A	N	R2X(совокупн.)
1	M1	OPLS-DA	1+7+0	563	0,944
2	M2	OPLS-DA	1+4+0	563	0,662
3	M3	PCA-X	11	563	0,977

График рассеяния показателей [M1]
 График рассеяния показателей [M2]
 График вклада [M2]
 График Хнабл [M2]
 График Хнабл [M1]

УФ_PCA_Simca_Импорт_Файл_25.M2 (OPLS-DA), Все партии 1-я Произ
 Масштабируется пропорционально R2X
 Окрашено в соответствии с Идент. набл. (Хорошая/Плохая)



R2X[1] = 0,142, R2Xo[1] = 0,186, Эллипс: Гостиничный комплекс T2 (95%)

Продолж. На
 ФИГ. 32-2

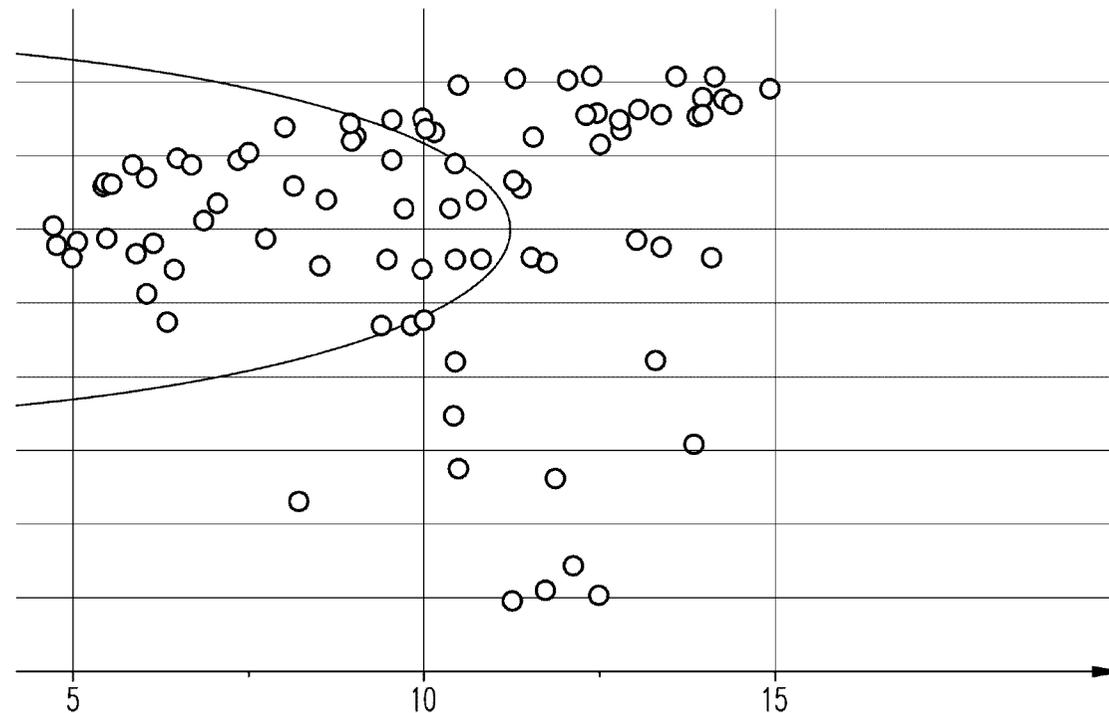
ФИГ. 32-1

55/67

R2Y(совокупн.)	Q2(совокупн.)	Дата	Заголовок	Иерархическая
0,799	0,784	5/3/2019	Все партии Без преобразования	▲
0,811	0,797	5/3/2019	Все партии 1-я произ	≡
	0,961	5/3/2019	Все партии РСА	▼

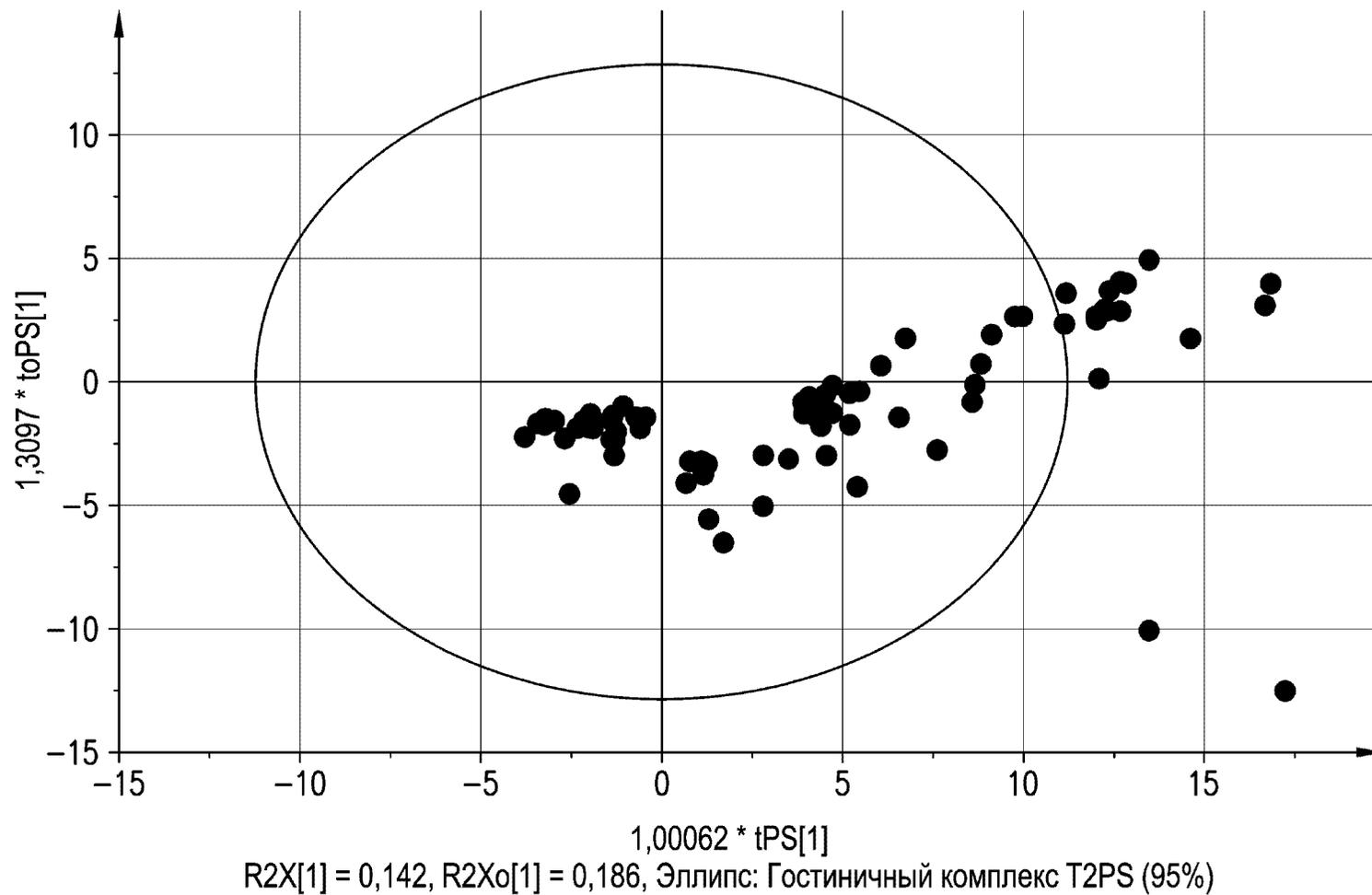
○ Плохая
● Хорошая

Продолж. От
ФИГ. 32-1



ФИГ. 32-2

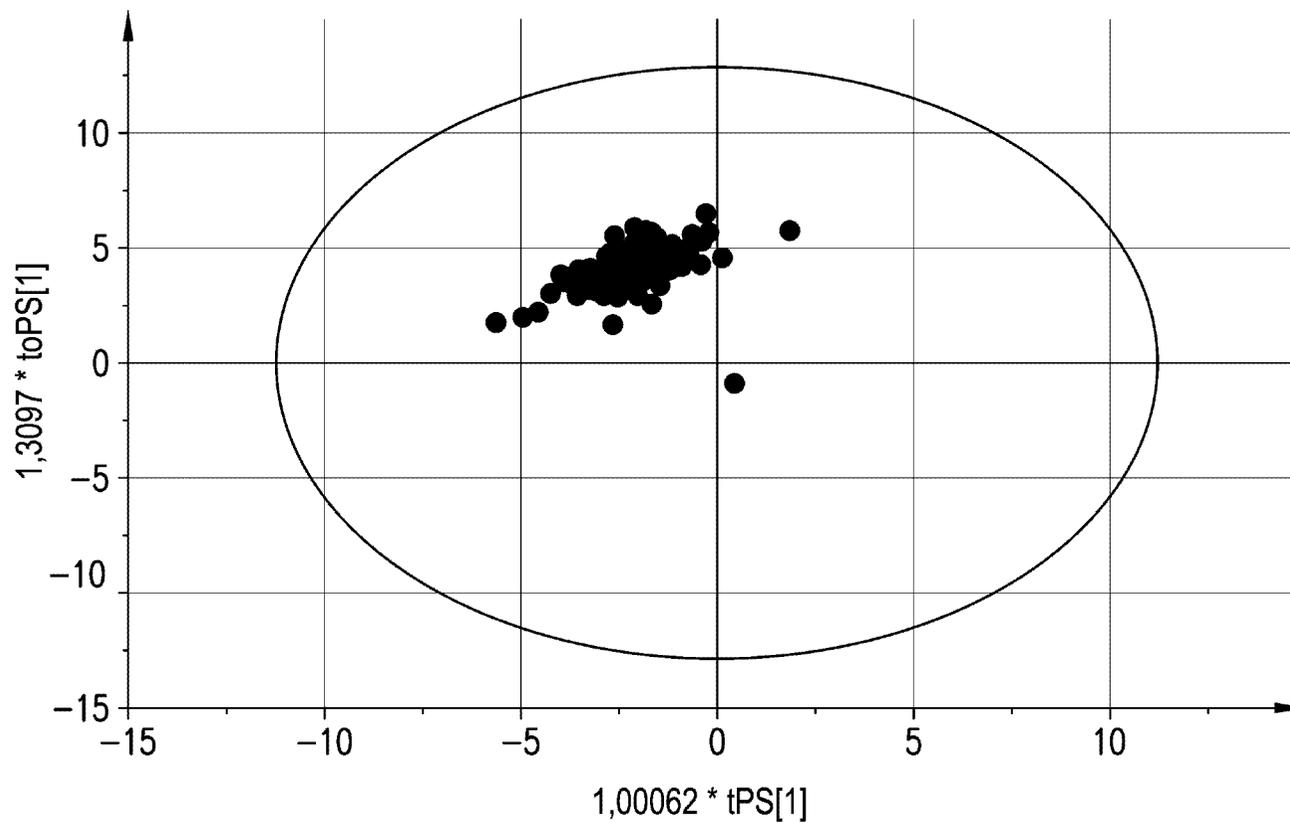
УФ_PCA_Simca_Импорт_Файл_25.M2 (OPLS-DA), Все партии 1-я Произ, PS-Матрица
X_2 Масштабируется пропорционально R2X



ФИГ. 33

57/67

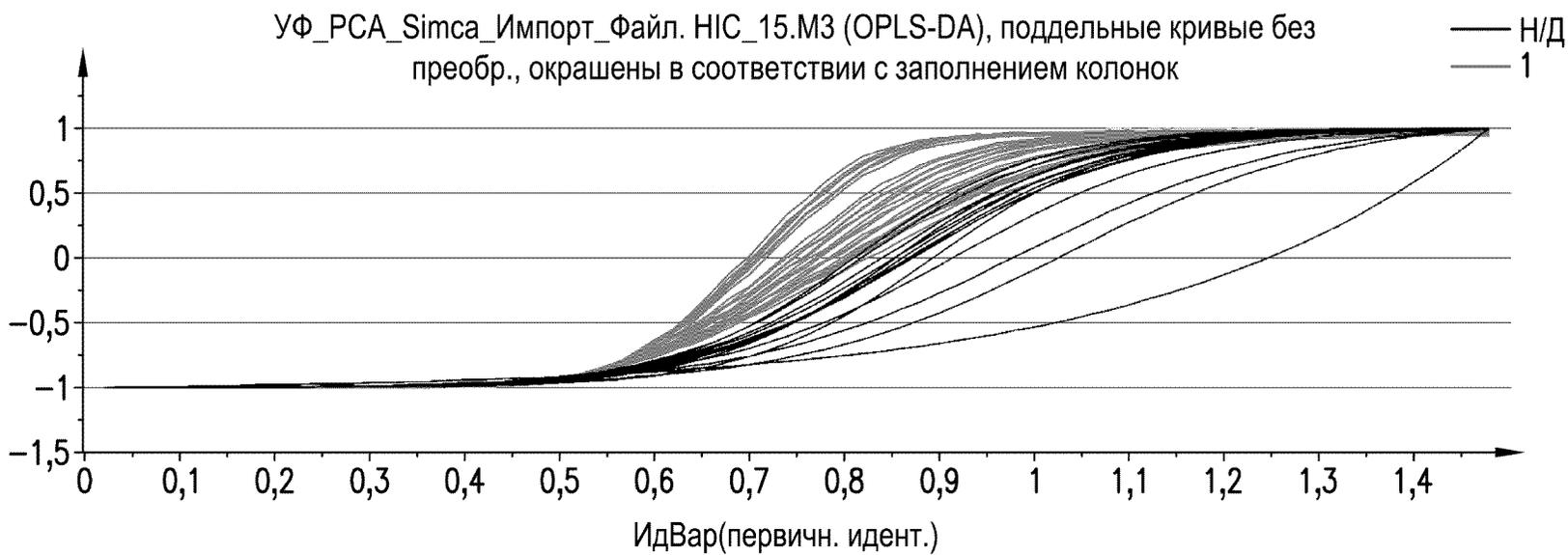
УФ_PCA_Simca_Импорт_Файл_25.M2 (OPLS-DA), Все партии 1-я Произ, PS-Матрица X_
Заполн18 Масштабируется пропорционально R2X



$R2X[1] = 0,142$, $R2Xo[1] = 0,186$, Эллипс: Гостиничный комплекс T2PS (95%)

ФИГ. 34

Хнабл(8183600011_1), Хнабл(8183600011_2),
Хнабл(8183600012_1), Хнабл(8183600012_2), ХО



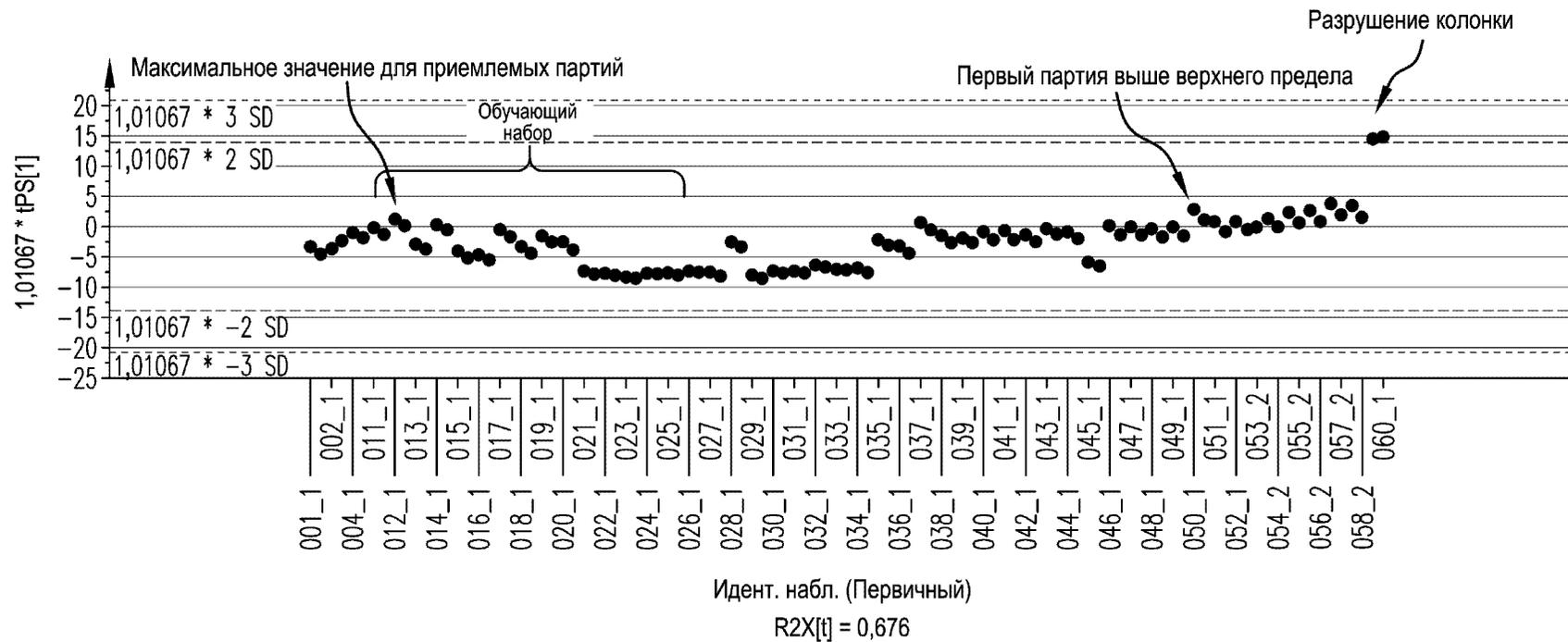
ФИГ. 35

УФ_PCA_Simca_Импорт_Файл_HIC_15.M9 (OPLS-DA), FC Партии 1-я произв
 Масштабируется пропорционально R2X Окрашено в соответствии с идент.
 набл. (Хорошая/Плохая)



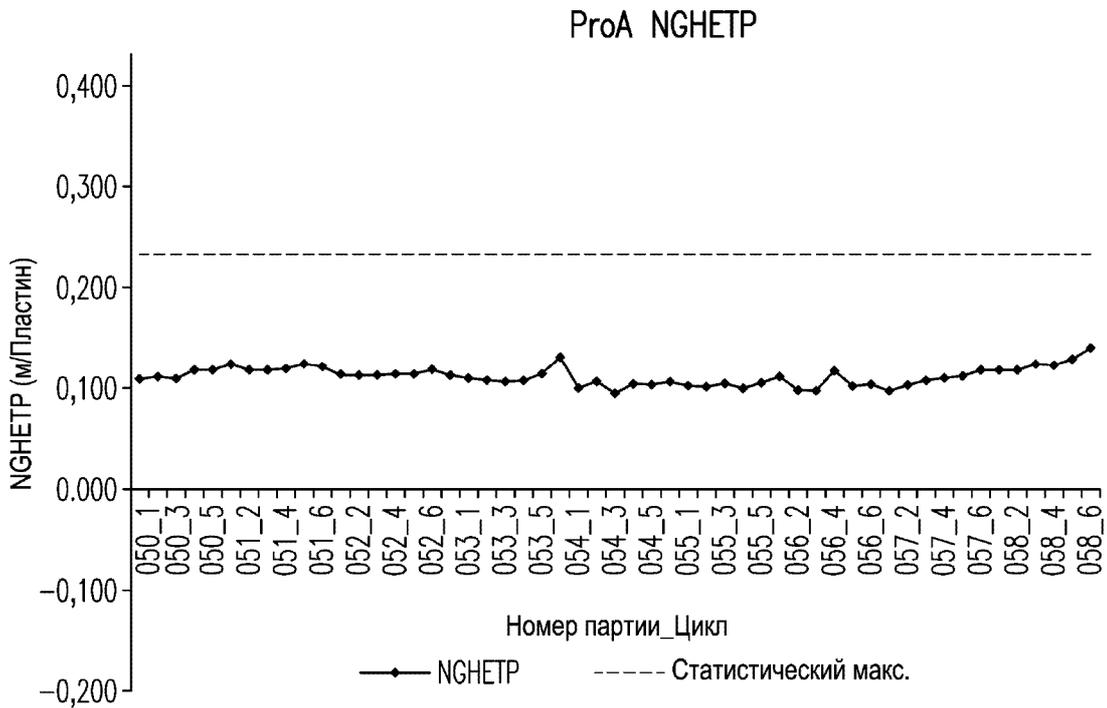
ФИГ. 36

УФ_PCA_Simca_Импорт_Файл, HIC_15.M9 (OPLS-DA), FC Партии 1-я прозв PS-Округл_заполн2
 Масштабируется пропорционально R2X Окрашено в соответствии с идент. набл.
 (Заполнение колонки)



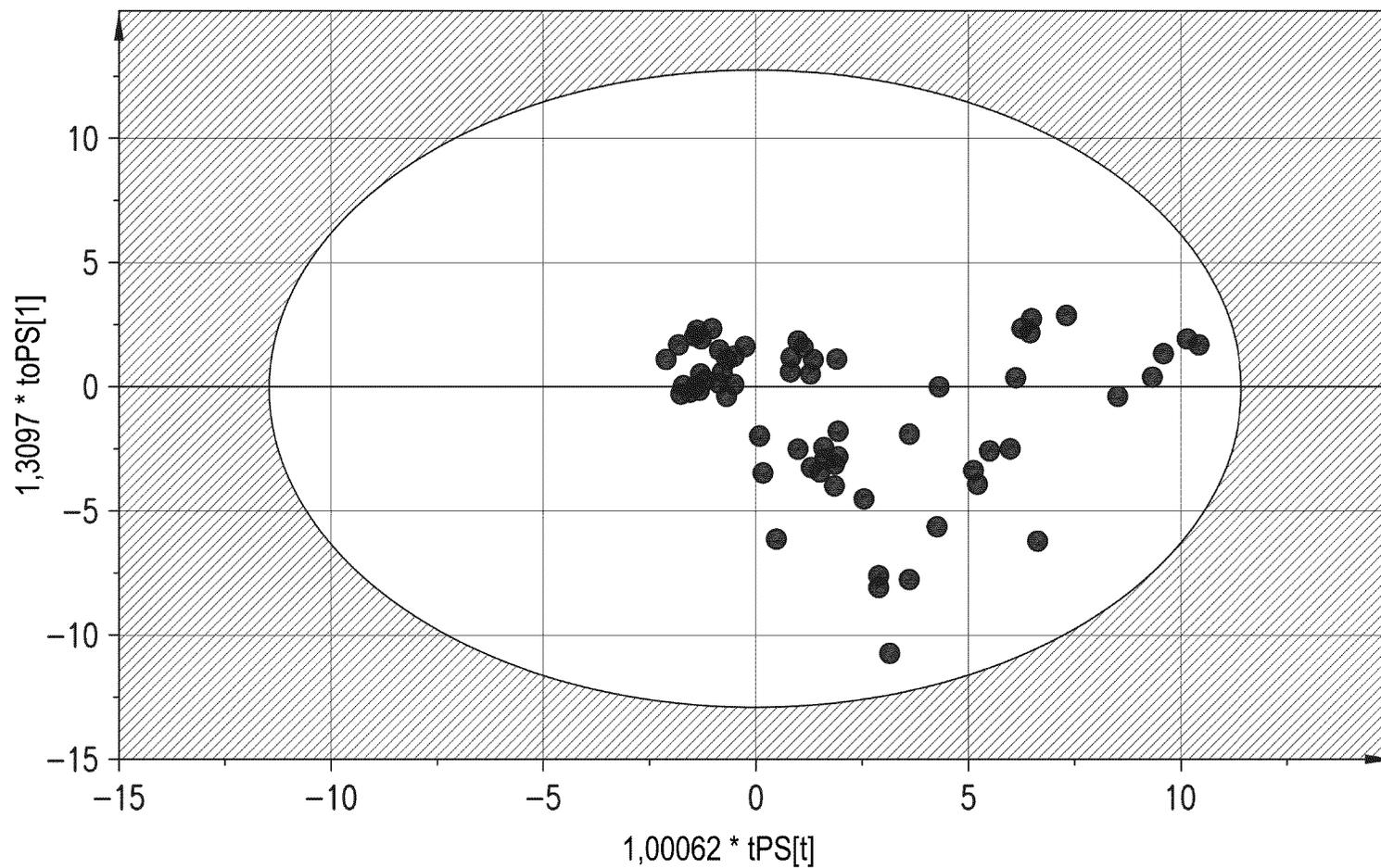
ФИГ. 37

62/67



ФИГ. 38А

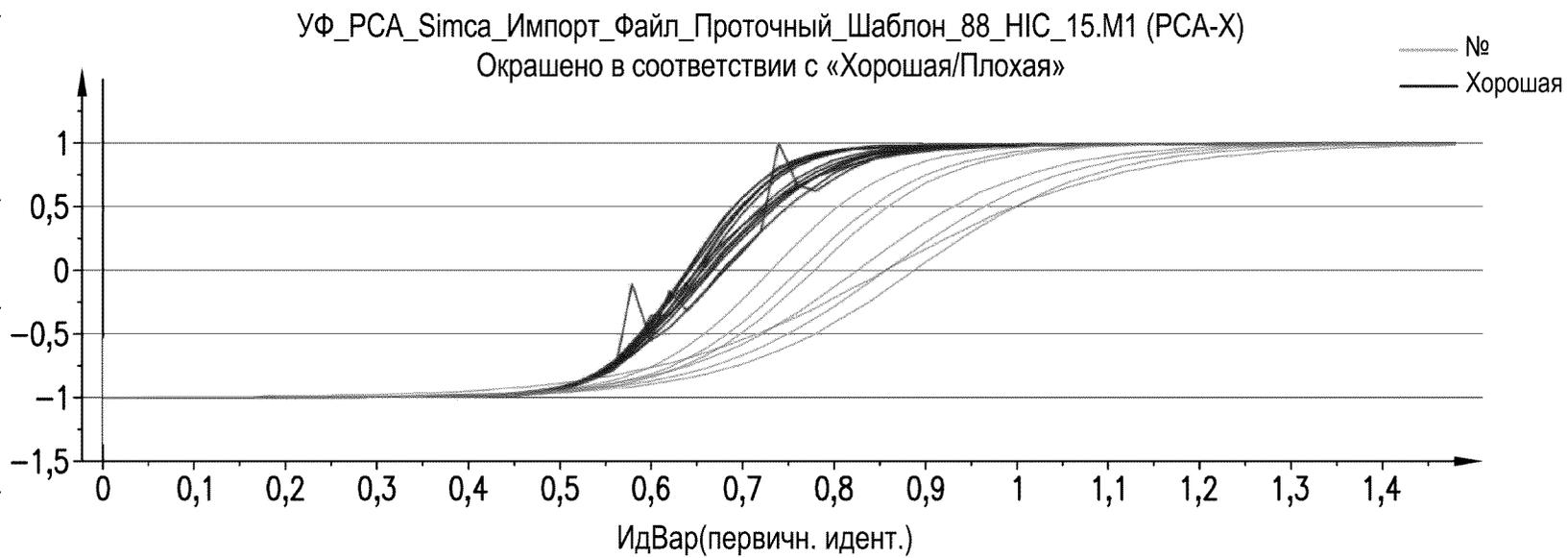
УФ_PCA_Simca_Импорт_Файл_25.M2 (OPLS-DA), Все партии 1-я Произ, PS-Матрица
X_Заполн11 Масштабируется пропорционально R2X



$R^2X[1] = 0,142$, $R^2Xo[1] = 0,186$, Эллипс: Гостиничный комплекс T2PS (95%)

ФИГ. 38В

Хнабл(8024000203_1), Хнабл(8024000204_1),
Хнабл(8024000205_1), Хнабл(8024000206_1), Х0



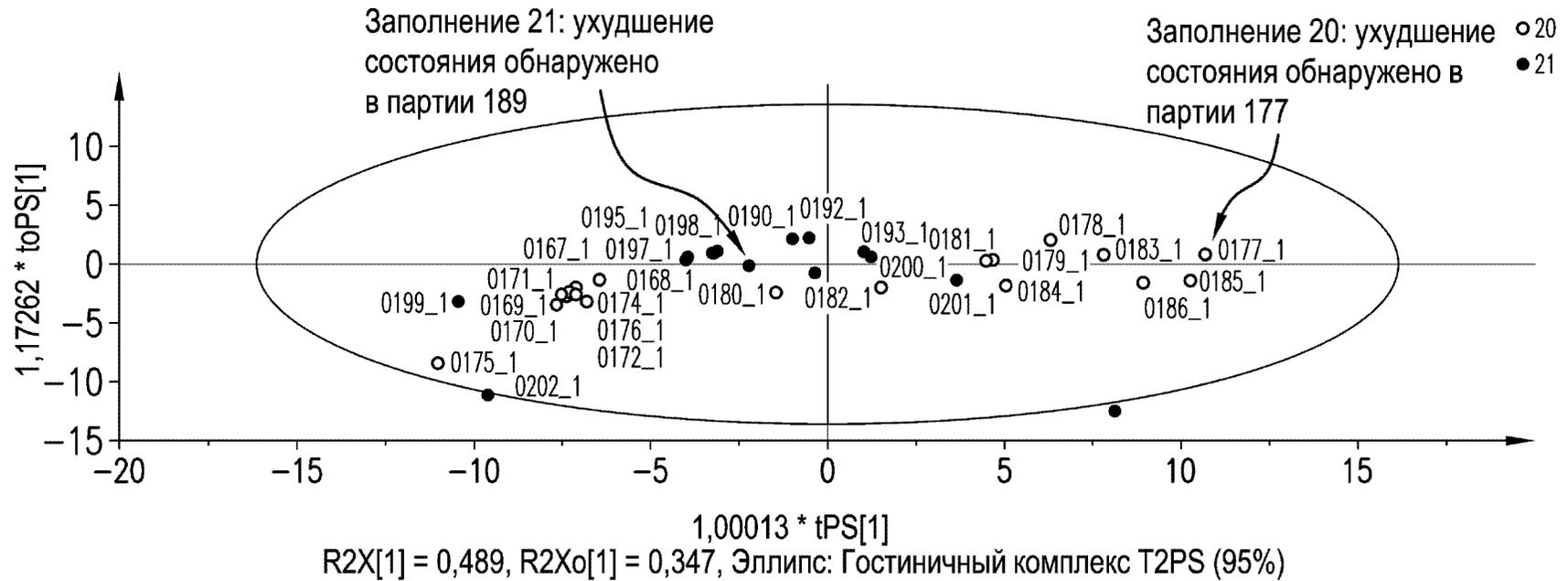
ФИГ. 39А

УФ_PCA_Simca_Импорт_Файл_Проточный_Шаблон_HIC_15.M3 (OPLS-DA)
Масштабируется пропорционально R2X Окрашивается в соответствии с классами в M3

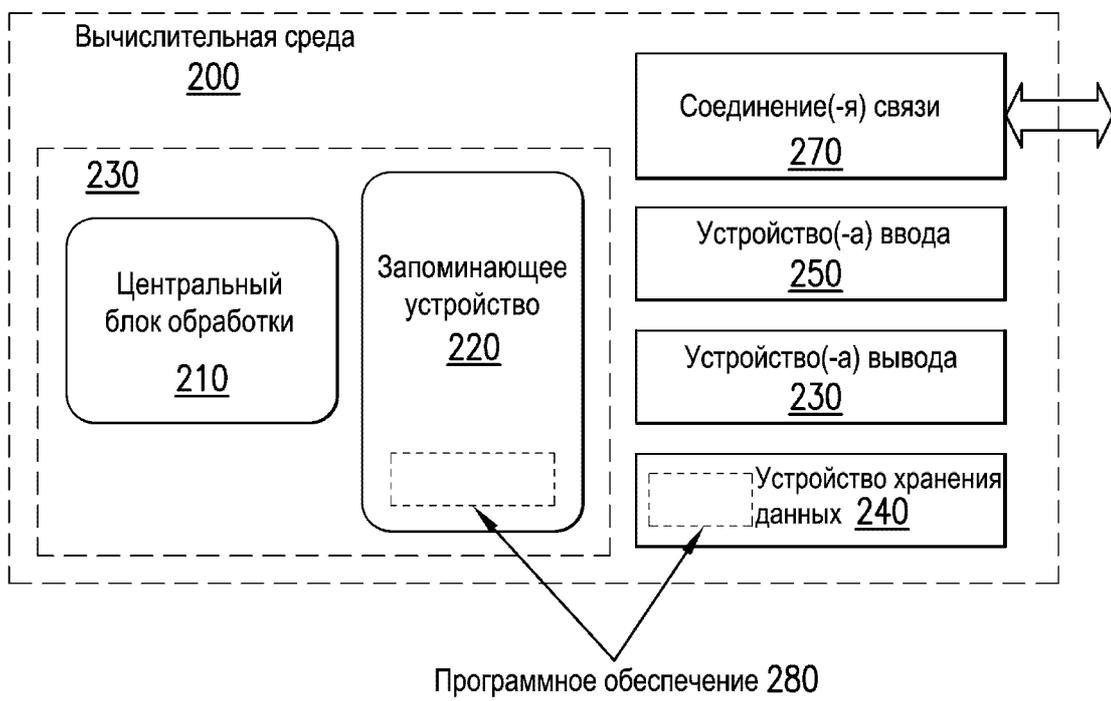


ФИГ. 39В

УФ_PCA_Simsa_Импорт_Файл_Проточный_Шаблон_88_Н1С_15.М3 (OPLS-DA),
PS-Округл_заполн21 Масштабируется пропорционально R2X Окрашено
в соответствии с идент. набл.



ФИГ. 40



ФИГ. 41