

УДК: 681.5.015

## Идентификация модели объекта при наличии неизвестных возмущений с широким частотным диапазоном на основе перехода к приращениям сигналов и отбора данных

М. Ю. Рябчиков<sup>а</sup>, Е. С. Рябчикова

Магнитогорский государственный технический университет им. Г. И. Носова,  
Россия, 455000, г. Магнитогорск, пр. Ленина, д. 38

E-mail: <sup>а</sup> mr\_mgn@mail.ru

*Получено 18.09.2023, после доработки — 14.11.2023.*

*Принято к публикации 20.11.2023.*

Работа посвящена проблеме создания модели со стационарными параметрами по ретроспективным данным в условиях неизвестных возмущений. Рассматривается случай, когда представительная выборка состояний объекта может быть сформирована с использованием ретроспективных данных, накопленных только в течение значительного интервала времени. При этом допускается, что неизвестные возмущения могут действовать в широком частотном диапазоне и могут иметь низкочастотные и трендовые составляющие. В такой ситуации включение в выборку данных разных временных периодов может привести к противоречиям и чрезвычайно снизить точность модели. В работе дан обзор подходов и способов согласования данных. При этом основное внимание уделено отбору данных. Дана оценка применимости различных вариантов отбора данных как инструмента снижения уровня неопределенности. Предложен метод идентификации модели объекта с самовыравниванием по данным, накопленным за значительный период времени в условиях неизвестных возмущений с широким частотным диапазоном. Метод ориентирован на создание модели со стационарными параметрами, не требующей периодической перенастройки под новые условия. Метод основан на совместном применении отбора данных и представлении данных отдельных периодов времени в виде приращений относительно начального для периода момента времени. Это позволяет уменьшить число параметров, которые характеризуют неизвестные возмущения при минимуме допущений, ограничивающих применение метода. В результате снижается размерность поисковой задачи и минимизируются вычислительные затраты, связанные с настройкой модели. Рассмотрены особенности применения метода при нелинейной модели. Метод использован при разработке модели закрытого охлаждения стали на агрегате непрерывного горячего оцинкования стальной полосы. Модель может использоваться при упреждающем управлении тепловыми процессами и при выборе скорости движения полосы. Показано, что метод делает возможным разработку модели тепловых процессов с секции закрытого охлаждения в условиях неизвестных возмущений, имеющих в том числе низкочастотные составляющие.

Ключевые слова: идентификация, большие данные, глобальная модель, приращения, неизвестные воздействия, отбор данных

Исследование выполнено за счет гранта Российского научного фонда № 23-29-10058 (<https://rscf.ru/project/23-29-10058/>).

UDC: 681.5.015

## Identification of an object model in the presence of unknown disturbances with a wide frequency range based on the transition to signal increments and data sampling

M. Yu. Ryabchikov<sup>a</sup>, E. S. Ryabchikova

Nosov Magnitogorsk State Technical University,  
38 Lenin ave., Magnitogorsk, 455000, Russia

E-mail: <sup>a</sup> mr\_mgn@mail.ru

*Received 18.09.2023, after completion – 14.11.2023.*

*Accepted for publication 20.11.2023.*

The work is devoted to the problem of creating a model with stationary parameters using historical data under conditions of unknown disturbances. The case is considered when a representative sample of object states can be formed using historical data accumulated only over a significant period of time. It is assumed that unknown disturbances can act in a wide frequency range and may have low-frequency and trend components. In such a situation, including data from different time periods in the sample can lead to inconsistencies and greatly reduce the accuracy of the model. The paper provides an overview of approaches and methods for data harmonization. In this case, the main attention is paid to data sampling. An assessment is made of the applicability of various data sampling options as a tool for reducing the level of uncertainty. We propose a method for identifying a self-leveling object model using data accumulated over a significant period of time under conditions of unknown disturbances with a wide frequency range. The method is focused on creating a model with stationary parameters that does not require periodic reconfiguration to new conditions. The method is based on the combined use of sampling and presentation of data from individual periods of time in the form of increments relative to the initial point in time for the period. This makes it possible to reduce the number of parameters that characterize unknown disturbances with a minimum of assumptions that limit the application of the method. As a result, the dimensionality of the search problem is reduced and the computational costs associated with setting up the model are minimized. It is possible to configure both linear and, in some cases, nonlinear models. The method was used to develop a model of closed cooling of steel on a unit for continuous hot-dip galvanizing of steel strip. The model can be used for predictive control of thermal processes and for selecting strip speed. It is shown that the method makes it possible to develop a model of thermal processes from a closed cooling section under conditions of unknown disturbances, including low-frequency components.

**Keywords:** identification, big data, global model, increments, unknown impacts, data sampling

*Citation:* *Computer Research and Modeling*, 2024, vol. 16, no. 2, pp. 315–337 (Russian).

The study was funded by a grant No. 23-29-10058 Russian Science Foundation, (<https://rscf.ru/project/23-29-10058/>).

## 1. Введение

В настоящее время сбор и хранение технологических данных является одной из неотъемлемых функций производства различных видов продукции. Ретроспективная информация, которую хранят накопленные данные, является ресурсом повышения эффективности производства. В [Drovandi et al., 2017] указывается три проблемы обработки больших данных: данные хранят информацию о процессе в условиях неизвестных возмущений; высокие вычислительные затраты; проблемы объединения данных разных источников. Модель может рассматриваться как инструмент решения проблем, присущих большим данным. Настроенная модель позволяет быстро решать задачи интерполяции и экстраполяции. Однако, кроме этого, синтез модели можно интерпретировать как выбор варианта разрешения проблем неоднородности и противоречивости ретроспективных данных.

Выбор структуры модели обычно производится с учетом свойств экспериментальных данных. Так, в [Elragal, Klischewski, 2017] вводится этап предварительной обработки данных, который предполагает выполнение действий, необходимых для решения проблем с качеством данных. При этом к этапу предварительной обработки данных авторы относят и выбор структуры модели. Отмечается, что при наличии больших данных исследование может начинаться с обработки данных для получения открытий, основанных непосредственно на данных. При этом оценки точности модели, полученные с применением ретроспективных данных, могут определять ее структуру. Основной проблемой открытий, основанных на данных, авторы называют эпистемологические ловушки, приводящие к чрезмерной роли эмпиризма.

В работе [Рябчиков, Рябчикова, 2022] рассмотрены противоречия подходов к моделированию, основанных на данных и на требованиях к модели. Указано, что недостатки модели со стационарными параметрами часто стремятся компенсировать с применением систем коррекции прогноза от модели с помощью дополнительной модели временного ряда ошибок или путем самонастройки основной модели. При этом могут использоваться различные модели временного ряда. В последние десятилетия существенное развитие получили модели, основанные на использовании фильтра Калмана. Так, например, в [Aved'yan, Tsyarkin, 1998] рассматривается задача прогноза в случае, когда возмущения могут быть представлены проинтегрированным белым шумом (броуновское движение). Предложенное решение эффективно при наличии у возмущений трендовых и низкочастотных компонентов.

Однако коррекция прогноза и самонастройка модели не являются универсальными средствами и могут приводить к существенным ошибкам при прогнозе. Причиной является то, что подобные меры могут привести к существенному непостоянству точности модели для разных состояний объекта. В таком случае, если состояние объекта изменяется быстро или даже скачкообразно, системы коррекции прогноза с применением дополнительной модели временного ряда или самонастройки основной модели могут оказаться неэффективны. Значительное время состояние объекта может отличаться от желаемого. Подробнее эта проблема рассмотрена нами в работах [Рябчиков, Рябчикова, 2022; Ryabchikov, Ryabchikova, 2022].

В то же время наличие модели со стационарными параметрами, которая адекватна задаче управления при ее тестировании по данным любого периода времени в прошлом, позволяет допустить, что модель и далее в будущем будет адекватна задаче управления. В работе [Bandara, Bergmeir, Hewamalage, 2021] подобные модели, которые настраиваются по данным временных рядов, полученных в разное время, названы глобальными. Авторы отмечают существенный рост интереса к созданию подобных моделей в последние годы.

Согласование больших данных, с целью получения адекватной задаче управления глобальной модели со стационарными параметрами, является задачей, для которой в настоящее время нет универсального решения. Так, авторы [Lemeshko, Lemeshko, Semenova, 2018] указывают на

то, что методы и критерии математической статистики неэффективны из-за высокой размерности больших данных. Поэтому предлагается оценивать группированные данные. В то же время группировка данных (усреднение, поиск медианы распределения и др.) может сделать прогноз чрезмерно неопределенным. Более того, при неопределенных свойствах неизвестных воздействий может быть неясно, к каким последствиям привела группировка данных. То есть сама оценка уровня определенности такого прогноза может быть затруднена. Вследствие этого весьма актуальным в настоящее время является развитие методов согласования данных.

Случай, когда неизвестные возмущения приводят к медленному изменению или дрейфу состояния процесса, является достаточно распространенным. Так, авторы работы [Wang et al., 2021] предложили систему прогноза деградации авиационных двигателей, которая предполагает нормирование данных индивидуально для различных режимов работы. Нормирование позволяет представить особенности деградации в виде сигналов, изменение которых имеет только монотонные и трендовые составляющие. В работах [Linkan, Nagi, 2014; Linkan, Nagi, Kharoufeh, 2015] предлагается использование накопленных и текущих данных для оценки параметров среды и определения остаточного срока службы компонентов процесса. Отмечается, что ключевой задачей является выявление схожих условий среды для агрегирования данных. Авторы [Wang, Liao, Ma, 2021] вводят зависимость вида функции деградации от условий среды.

Сложность расширения сферы применения предложений [Wang et al., 2021; Linkan, Nagi, 2014; Linkan, Nagi, Kharoufeh, 2015; Wang, Liao, Ma, 2021] заключается в том, что заранее может быть неизвестно, какие характеристики, сигналы и параметры системы подвержены неизвестным возмущениям. Также могут быть неизвестны особенности изменения свойств неизвестных возмущений во времени. Это приводит к неопределенности решения задачи оценки неизвестных возмущений. Например, авторы [Hajiha, Liu, Hong, 2021] рассматривают конкурирующие процессы деградации вследствие влияния внешней среды и определяют вероятностные оценки остаточного срока службы. Представляет интерес работа [Pannocchia, Rawlings, 2003], где показано, что для упреждающего управления без смещения может быть недостаточным учет числа интегрируемых возмущений, равного числу управляемых переменных. Достаточное число интегрируемых возмущений в общем случае равно числу измеряемых переменных. Также авторы отмечают целесообразность определения того, какие возмущения имеют тренд, с последующей разработкой моделей именно для таких возмущений.

Однако возможна ситуация, когда не удастся разработать вариант модели, которая позволяет согласовать данные разных временных периодов путем представления возмущений в виде сигналов, имеющих тренд. Иногда не удастся создать модель для представления возмущений в виде сигналов, статистические свойства которых возможно прогнозировать. Наиболее проблемным является вариант, когда данные разных периодов принципиально не удастся согласовать. На практике это проявляется в том, что для каждого отдельного периода времени модель имеет свои оптимальные значения параметров настройки, не применимые для других периодов времени. Это затрудняет разработку глобальной модели для прогнозирования, которую можно было бы использовать без перенастройки. Одним из негативных последствий этого является то, что затруднительно сформировать представительную выборку различных состояний объекта для получения адекватной в статистическом плане модели.

Целью данной работы является разработка метода идентификации модели со стационарными параметрами, в случае когда получение представительной выборки требует использования данных за значительный период времени. При этом допускается, что данные получены в условиях неизвестных возмущений с широким частотным диапазоном, которые могут иметь низко-частотные и трендовые составляющие. Модель должна быть адекватна задаче управления при тестировании по ретроспективным данным любого периода времени.

## 2. Большие данные при синтезе систем управления

К настоящему времени давно сформулированы отличительные аспекты больших данных. Большие данные отличаются большими объемами, разнообразием формы представления, высокой скоростью поступления. Однако по-прежнему в полной мере не ясно, какую выгоду из больших данных можно извлечь при синтезе систем управления.

В работе [Ryabchikov, Ryabchikova, 2022] нами рассмотрен аспект больших данных, связанный с тем, что данные, собранные за значительный временной период, потенциально могут использоваться для получения интегральных оценок эффективности разных вариантов управления, включая варианты управления, ранее не опробованные на практике. Это предполагает получение моделей, демонстрирующих высокий уровень постоянства точности или прогнозируемости точности в зависимости от состояний сред, с которыми взаимодействует моделируемый объект. Причиной является возможность получения принципиально новых состояний сред и объекта при использовании неопробованных ранее на практике вариантов управления. Большие данные являются средством оценки степени постоянства или вида и параметров зависимостей точности моделей от состояний сред. Демонстрация моделью точности, оценки которой являются прогнозируемыми (желательно постоянными) в течение значительного периода времени, который охватывают большие данные, является аргументом в пользу выбора той или иной структуры и значений стационарных параметров модели.

Для высокой эффективности систем управления прогнозирующую модель со стационарными параметрами дополняют моделью временного ряда возмущений. Распространенный подход к созданию модели временного ряда предполагает изучение статистических свойств моделируемой величины и выбор типа модели с учетом свойств. Подход сложился достаточно давно и предполагает определенную последовательность идентификации в узком и широком смысле. Так, в работе [Райбман, 1979] задача идентификации, связанная только с оценкой параметров по данным «вход – выход», относится к задаче идентификации в узком смысле. При решении задачи требуется предварительно определить некоторые общие характеристики, то есть решить задачу в широком смысле.

В работе [Рябчиков, Рябчикова, 2022] при сравнении подходов к моделированию, основанных на данных и на требованиях, нами обсуждается новая задача. Предлагается использовать наиболее простые модели временного ряда, а выбор структуры и настроек модели со стационарными параметрами осуществлять с позиции эффективности выбранного вида модели временного ряда. Тогда с учетом предложений в [Ryabchikov, Ryabchikova, 2022] задача синтеза системы упреждающего управления в рамках подхода, основанного на данных, сводится к синтезу модели со стационарными параметрами, демонстрирующей прогнозируемые оценки точности в течение любого интервала времени, который охватывают большие данные. В таком случае задача настройки модели тесно связана с задачей согласования данных.

## 3. Согласование данных при настройке модели

При согласовании данных можно выделить два различных подхода. Первый подход предполагает обработку данных путем их агрегирования или отбора. Второй подход основан на совместном поиске настроек модели и сигнала неизвестных возмущений. Рассмотрим особенности подходов.

Для согласования данных, полученных в различные периоды времени, в рамках первого подхода часто используют отбор данных. При этом, как правило, преследуют две основные цели: устранение дисбаланса в выборке данных различных классов и устранение влияния неизвестных возмущений.

Дисбаланс классов приводит к гетероскедастичности остатков и непостоянству точности модели для разных ситуаций. Для обеспечения баланса классов известна обширная группа методов сэмплинга [Рябчиков, Рябчикова, Сунаргулова, 2016]. Однако их применение, как правило, предполагает отбор наборов данных из всей совокупности информации. При этом значения откликов для схожих условий могут иметь существенный разброс вследствие неизвестных возмущений, и точность полученных моделей будет относительно низкой. Причиной являются низкочастотные и трендовые составляющие неизвестных возмущений в ретроспективных данных, полученных за значительный временной период. Применение методов сэмплинга часто дополняют агрегированием данных (среднее, медиана распределения и др.) для минимизации смещения при прогнозе. Более развитые методы могут использовать статистические гипотезы. Примером является предложенная в [Feng et al., 2018] система для определения источника возмущений в системах энергоснабжения на основе метода «случайный лес» (Random forest). Система предполагает случайный выбор множества вариантов входных переменных (из общей их совокупности), для которых индивидуально формируются выборки (применяется алгоритм SMOTE — отбор данных) с последующим построением дерева решений и их организацией в лес решений. Далее определение типа источника возмущений выполняется путем голосования между результатами, полученными при использовании разных вариантов обработки данных. Полученные результаты оцениваются методом out-of-bag, что предполагает сравнение ошибок обучения классификаторов. Конструирование леса повторяется до получения приемлемой ошибки. Схожий метод был применен авторами [Maxwell, Rajabi, Esterle, 2019] для классификации метаморфизованного угля по геофизическим данным. Предложения [Feng et al., 2018; Maxwell, Rajabi, Esterle, 2019] ориентированы на ситуацию избыточного числа переменных, характеризующих состояние процесса. В то же время в случае, если состояние процесса характеризуют относительно небольшое число технологических параметров, то, вероятно, все они будут использованы при создании модели. Случайность, положенная в основу методов, затрудняет гарантии того, что ошибка будет снижена для приемлемого для практических решений уровня. Общей проблемой рассмотренных методов является то, что они не позволяют существенно уменьшить доверительные интервалы. Вследствие этого важной целью отбора данных становится устранение влияния неизвестных возмущений.

При использовании отбора данных для устранения влияния неизвестных возмущений принимают гипотезы. Одной из распространенных гипотез является допущение о том, что неизвестные возмущения обусловлены различными видами продукции или режимами работы. Например, в работе [Kozjek et al., 2019] предложена схема автоматизированной обработки данных с целью диагностики проблем при литье пластмасс под давлением. Отбор данных для настройки модели предполагает три основных этапа: выбор данных, связанных с производством продукта определенного типа; устранение данных периодов перехода между состояниями статуса; удаление части данных мажоритарного класса (undersampling) о работе при нормальном режиме. Подобный отбор предполагает, что свойства возмущений связаны с производством продукта определенного типа. Можно указать два недостатка метода. Во-первых, данных по некоторым видам продукции может оказаться недостаточно для создания отдельной модели. Во-вторых, при запуске в производство нового вида продукции потребуется предварительно накопить данные. А ведь именно модель может являться тем инструментом, который может снизить затраты на перенастройку производства.

Представляет интерес метод отбора данных, предложенный в [Kazarinov, Khasanov, 2019; Казаринов, Хасанов, 2019]. Предлагается интерактивный метод нейроруправления контуром измельчения шихты при производстве цемента для полуавтоматических мельниц. Для настройки модели авторами предложен алгоритм, ориентированный на выбор максимально большого подмножества данных процесса, для которого выполняется принятое условие непротиворечивости.

В качестве условия непротиворечивости принято попадание сигналов, характеризующих состояние процесса, в диапазоны значений. Причем выбор диапазонов осуществляется совместно с отбором данных. Можно указать два недостатка предложенного метода. Соответствие переменных состояния процесса диапазонам значений не означает отсутствия неизвестных возмущений. Неясно, всегда ли можно использовать модель, или только в отдельные периоды времени, или только при отдельных состояниях процесса. То есть отбор данных некоторых периодов во времени должен, по-видимому, иметь некоторые обоснования.

В работе [Shardt, Brooks, 2018] предложен алгоритм сегментации данных процесса в цинковой флотационной камере, основанный на применении многочленов Лаггера [Wahlberg, 1991; Wahlberg, 1994]. Ортогональность полиномов Лаггера позволяет легко удалять ненужные компоненты модели без воздействия на остальные параметры. Алгоритм предполагает перебор данных согласно времени их поступления в базу. После добавления в сегмент нового набора данных оцениваются условия корректности аппроксимации и принимается решение либо о продолжении формирования текущего сегмента, либо о начале нового сегмента данных. Авторами [Shardt, Brooks, 2018] показано, что решение может быть получено только при правильном выборе переменных и исключении переменных, в действительности не оказывающих влияние. Предложенный алгоритм предполагает, что значения некоторых из параметров модели изменяются во времени, оставаясь тем не менее постоянными для отдельных сегментов. Проблемой применения алгоритма в общем случае является то, что данных отдельного сегмента может оказаться недостаточно для настройки модели. При этом поиск вариантов объединения данных множества сегментов затруднен объемами данных и большим числом сегментов. Также применение полиномов Лаггера может быть затруднено существенной нелинейностью и разрывами зависимостей, на которых основана модель.

Второй подход к согласованию данных при настройке модели ориентирован на объяснение противоречий в данных неизвестными возмущениями. Для их определения разрабатывается наблюдатель возмущений. Подробный обзор методов идентификации модели с определением неизвестных возмущений выполнен в [Andrievsky, Furtat, 2020]. Обзор охватывает историю развития многих методов. Распространенной является группа методов [Francis, Wonham, 1976; Краснова, 2018; Malikov, 2018] определения неизвестных внешних воздействий, предполагающих создание модели, основанной на дифференциальных уравнениях. Решения уравнений описывают внешние воздействия, а искомые неизвестные начальные условия характеризуют неопределенные свойства воздействий. Примечательно в обзоре [Andrievsky, Furtat, 2020] то, что, несмотря на подробное описание развития и классификацию методов определения неизвестных возмущений, основанных на использовании модели, проблемы выбора вида модели и ее настройки не упоминаются. При синтезе методов часто изначально предполагается, что возможно создать модель, которая удовлетворяет требованиям. Однако создание модели требует подготовки данных. Применить методы [Francis, Wonham, 1976; Краснова, 2018; Malikov, 2018] при работе с большими данными без их дополнительной обработки затруднительно. Причинами этого являются как высокая размерность, так и большое число наборов данных во временных последовательностях. Неясно, как организовать определение неизвестных начальных условий совместно с обработкой данных.

Основой подобной обработки данных может являться другой широко применяемый на практике способ, основанный на принятии гипотезы о том, что неизвестные возмущения оказывают малое влияние на динамику процесса. Причиной, которая позволяет принять гипотезу для некоторых интервалов времени, являются характерные особенности известных возмущений. Например, значительная амплитуда ступенчатого воздействия при получении переходной характеристики позволяет пренебречь влиянием неизвестных возмущений. Так, в работе [Ryabchikov, 2020] принято допущение о выполнении принципа суперпозиции применительно к действию

возмущений на выходную переменную. Для определения траектории переходной характеристики настраивается модель временного ряда. Настройка производится по данным до возмущения (изменение управляющего воздействия). Далее из измеренных после возмущения данных вычитается прогноз (экстраполяция) состояния процесса, полученный с применением настроенной модели временного ряда. Недостатком метода является то, что он не позволяет устранить влияние на процесс низкочастотных компонентов неизвестных возмущений. Причиной этого является то, что модель временного ряда создается по данным относительно небольшого периода времени. Это затрудняет согласование данных, между получением которых прошло значительное время. Целесообразно объединить на основе единого метода преимущества отбора данных и совместного поиска настроек модели и сигнала неизвестных возмущений. Это требует принятия ряда ограничений и изучения влияния отбора данных на задачу определения неизвестных возмущений.

#### 4. Определение проблемы

Схема на рис. 1 определяет рассматриваемый в данной работе случай, когда синтез модели затруднен. Согласно схеме наиболее проблематичным является случай при одновременном наличии ряда обстоятельств:

- 1) для получения представительной выборки необходимы данные за значительный период времени;
- 2) требуется разработка глобальной модели со стационарными параметрами;
- 3) неизвестные возмущения имеют низкочастотные компоненты;
- 4) неизвестные возмущения имеют не только низкочастотные компоненты.

Проблемы нет, если для создания модели достаточно данных, полученных за относительно небольшой интервал времени. В таком случае могут использоваться сложные модели временного ряда, определяющие неизвестные возмущения, которые настраиваются совместно с моделью объекта. При отборе данных применимы методы сегментации данных, основанные на их непротиворечивости (например, [Kazarinov, Khasanov, 2019; Казаринов, Хасанов, 2019; Shardt, Brooks, 2018]).

Проблемы также нет, если разработка глобальной модели со стационарными параметрами принципиально не требуется. Это возможно, например, если состояние объекта изменяется медленно и возможен синтез такого управления по отклонению от заданного состояния, которое не приводит к недопустимым по величине или по продолжительности отклонениям. В таком случае наличие модели не является принципиально необходимым и модель является средством повышения качества управления. Такая модель может постоянно подстраиваться под текущий режим работы объекта. При изменении режима качество управления гарантируется за счет управления по отклонению.

Если неизвестные возмущения не имеют низкочастотных компонентов, то агрегирование данных может использоваться для устранения влияния прочих компонентов неизвестных возмущений. При этом в некоторых случаях при получении модели статических свойств возможно игнорирование наличия у объекта динамических свойств.

В случае когда неизвестные возмущения имеют только низкочастотные составляющие, то применимы предложения, где предлагается синтез единой модели временного ряда неизвестных возмущений для всего периода времени, который охватывают большие данные. Синтез модели осуществляется совместно с глобальной моделью процесса. Подобные решения предложены,



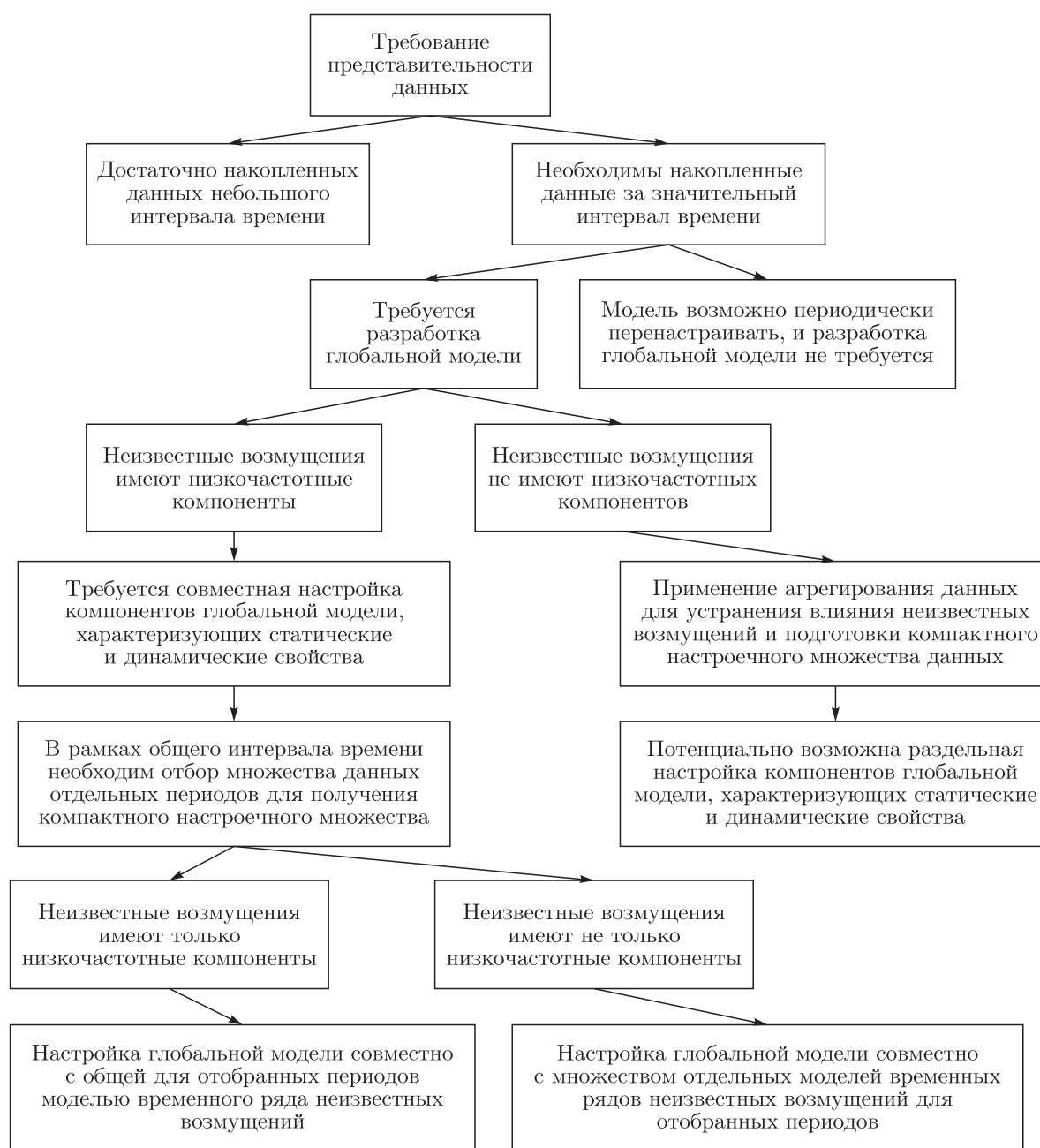


Рис. 1. Ситуации с различными требованиями к объему ретроспективных данных и различными свойствами неизвестных возмущений

например, в [Wang et al., 2021; Linkan, Nagi, 2014; Linkan, Nagi, Kharoufeh, 2015; Wang, Liao, Ma, 2021]. При неопределенности решения изучают возможные последствия разных вариантов.

Данная работа посвящена случаю, когда неизвестные возмущения имеют не только низкочастотные составляющие. В такой ситуации синтез глобальной модели со стационарными параметрами затруднительно осуществить совместно с моделью временного ряда неизвестных возмущений. Если использовать простые модели временного ряда, то доверительные интервалы для прогнозов от глобальной модели будут чрезмерно велики. Причиной является то, что простые модели временного ряда не позволят учесть составляющие неизвестных возмущений, отличные от низкочастотных. В то же время усложнение моделей временного ряда ограничено снижением

их адекватности и ростом вычислительной нагрузки, что связано с высокой степенью нелинейности и ростом числа настраиваемых параметров.

Одним из возможных направлений решения проблемы является замена единой сложной модели временного ряда набором простых моделей. Каждая из таких моделей создается по данным отдельного небольшого интервала времени (относительно общего периода времени, который охватывают большие данные). Это требует отбора данных.

## 5. Аспекты влияния отбора данных на идентификацию модели

Из обзора следуют две основные цели отбора данных: снижение гетероскедастичности остатков, устранение влияния неизвестных возмущений на смещение прогноза от модели. Устранение смещения прогноза достигается двумя основными способами.

Первый способ предполагает отбор данных таких периодов времени, для которых отсутствуют существенные противоречия [Kazarinov, Khasanov, 2019; Казаринов, Хасанов, 2019; Wahlberg, 1991; Wahlberg, 1994]. Способ позволяет исключить влияние низкочастотных составляющих неизвестных возмущений. Поскольку при отборе данных нет ограничений на уровень управляющих воздействий, то могут быть отобраны значительные объемы данных. При этом следует ожидать, что влияние составляющих неизвестных возмущений на более высоких частотах будут исключено путем агрегирования данных. Следует ожидать существенного снижения эффективности методов агрегирования с ростом размерности исходных данных. При значительном числе входных сигналов модели даже больших объемов данных может оказаться недостаточно. Кроме того, потенциальная польза модели, созданной с применением такого отбора данных, может быть невысока, поскольку неясно, в какие периоды времени можно использовать прогноз, а в какие — нет.

Второй способ предполагает отбор данных периодов времени, когда уровень управляющих воздействий велик [Ryabchikov, 2020]. Данный способ, напротив, для многих объектов будет эффективен при устранении влияния составляющих неизвестных возмущений, отличных от низкочастотных составляющих. Это объясняется тем, что управляющие воздействия для многих объектов сосредоточены в диапазоне, отличном от низкочастотного диапазона. Поскольку данных периодов времени, где уровень управляющих воздействий велик, может быть относительно мало, то эффективность применения агрегирования данных снизится. Следует ожидать, что при втором способе отбора данных низкочастотные составляющие неизвестных возмущений будут оказывать сильное влияние на смещение прогноза от модели.

Устранить недостаток второго способа отбора данных возможно путем разработки наблюдателя низкочастотных составляющих неизвестных возмущений. При использовании типового подхода неопределенные свойства возмущений находят при решении уравнений путем определения неизвестных начальных условий. Однако отбор данных приводит к ситуации, когда для каждого отдельного периода времени существуют свои начальные условия. Если отбираются данные значительного числа интервалов времени, то размерность задачи поиска параметров модели оказывается чрезвычайно высокой.

## 6. Принятые допущения относительно свойств объекта и неизвестных воздействий

Создание глобальной модели целесообразно для объекта, обладающего свойством самовыравнивания. Объект с самовыравниванием потенциально допускает создание модели на основе таких зависимостей между входными и выходными переменными в установившемся режиме, что все или некоторые параметры этих зависимостей стационарны. При управлении подобные модели могут использоваться в случае быстрого изменения состояния объекта, когда управление

по отклонению потенциально неэффективно. Также подобные модели имеют сферу применения, связанную с прогнозом оценок эффективности процессов и производства в целом с участием моделируемого объекта. Таким образом, допускаем, что объект обладает свойством самовыравнивания (допущение 1).

Высокая размерность задачи настройки модели, связанная с отбором данных множества отдельных временных последовательностей, приводит к целесообразности минимизации числа параметров, которые характеризуют неизвестные начальные условия. Поэтому целесообразно представить действие неизвестных возмущений одним параметром, имеющим индивидуальное значение для каждого отдельного сегмента данных. Поскольку модель может иметь множество входов, то минимизацию числа неизвестных возмущений обеспечивает структурное представление неизвестных воздействий как смещение значения выхода модели.

Как в [Ryabchikov, 2020], допустим, что применительно к действию возмущений на выходную переменную выполняется принцип суперпозиции (допущение 2). Тогда динамику выхода модели можно объяснить двумя компонентами суммы, а именно компонентами, связанными с действием известных и неизвестных возмущений. Для того чтобы компонент, связанный с действием неизвестных возмущений, потенциально мог находиться на постоянном уровне, объект должен обладать свойством самовыравнивания. Однако в случае, когда неизвестные возмущения действуют в широком частотном диапазоне, трудно гарантировать, что компонент, связанный с их действием, установился и неизменен в течение всего периода времени, который соответствует отдельному отобранному сегменту данных. Вместо этого допустим, что на интервале времени отдельного сегмента данных действие компонентов неизвестных возмущений в диапазоне частот более низком, чем диапазон частот известных возмущений, проявляется в изменении состояния процесса с постоянной скоростью (допущение 3). Применительно к компонентам неизвестных возмущений, действующим в диапазоне частот известных возмущений, примем, что влияние неизвестных возмущений пренебрежимо мало по сравнению с влиянием известных возмущений (допущение 4).

## 7. Предпосылки к отбору данных

Выбор интервала времени, в течение которого неизвестные возмущения предположительно имеют преимущественно низкочастотные составляющие, можно осуществить путем оценки динамики изменения выходных и входных переменных объекта. Если объект находится в состоянии равновесия и при этом значения его входных и выходных переменных находятся на постоянном уровне, то можно допустить, что неизвестные возмущения имеют только низкочастотные составляющие. Отбор данных с учетом указанной предпосылки требует выбора критерия оценки стабильности и предельных значений критерия. Если требуется увеличить объем отобранных данных, то необходимо снизить требования к стабильности значений входных и выходных переменных объекта. При этом гарантировать, что неизвестные возмущения имеют только низкочастотные составляющие, может быть затруднительно, и нужны иные предпосылки к отбору данных.

Такой предпосылкой может являться допущение о том, что результат действия неизвестных возмущений пренебрежимо мал по отношению к результату действия известных возмущений. Отбор данных с учетом указанной предпосылки может осуществляться путем оценки особенностей изменения известных возмущений. Например, критерий может оценивать скорость изменения и амплитуду известных воздействий.

Между двумя указанными предпосылками существует противоречие. Первая предпосылка требует стабильности входных сигналов, а вторая — их интенсивного изменения. Указанное противоречие можно разрешить путем поиска двух таких последовательно расположенных интервалов времени, что в течение первого интервала выполняется первая предпосылка, а в течение

второго — вторая. При этом допускаем, что в течение второго интервала неизвестные возмущения по-прежнему представлены только низкочастотными компонентами. Это позволяет использовать данные второго интервала для последующей настройки модели.

## 8. Свойство самовыравнивания и результативность отбора данных

Наличие у объекта свойства самовыравнивания приводит к тому, что после возмущения объект приходит к равновесному состоянию и без воздействий от регулятора. Осуществить управление подобным объектом в ручном режиме существенно проще, чем объектом без самовыравнивания. Для объекта без самовыравнивания действие возмущений может привести к нарушению равновесия и началу непрерывного изменения состояния объекта. В случае объекта с самовыравниванием, несмотря на действие возмущений, состояние объекта может остаться в пределах допустимого. При этом значения входных переменных объекта продолжительное время могут сохраняться на постоянных уровнях. Вследствие этого для объекта с самовыравниванием отбор данных, который осуществляется на основе принятых предпосылок, может быть более результативен в плане числа отобранных периодов времени при равных ограничениях на отбор.

## 9. Метод идентификации по данным в виде приращений относительно начальных для сегмента времени значений

Разработка метода требует определения структуры модели объекта с самовыравниванием и конкретизации проблем настройки моделей с такой структурой. Необходимо предложить вариант представления сигналов модели, а также набор условий для отбора настроечных данных, которые были бы ориентированы на создание глобальной модели, обеспечивающей несмещенный прогноз при минимизации вычислительных затрат на настройку.

### 9.1. Представление модели объекта с самовыравниванием

В качестве модели линейной системы со стационарными параметрами, имеющей один вход и выход, а также обладающей свойством самовыравнивания, примем типовое дифференциальное уравнение вида

$$\begin{aligned} a_n \frac{d^n y(\tau)}{d\tau^n} + a_{n-1} \frac{d^{n-1} y(\tau)}{d\tau^{n-1}} + \dots + a_1 \frac{dy(\tau)}{d\tau} + y(\tau) = \\ = b_m \frac{d^m x(\tau)}{d\tau^m} + b_{m-1} \frac{d^{m-1} x(\tau)}{d\tau^{m-1}} + \dots + b_1 \frac{dx(\tau)}{d\tau} + b_0 x(\tau), \end{aligned} \quad (1)$$

где  $x$ ,  $y$  — соответственно входной и выходной сигналы.

Введем обозначения:

$$A(\tau) = \left( a_n \frac{d^n y(\tau)}{d\tau^n} + a_{n-1} \frac{d^{n-1} y(\tau)}{d\tau^{n-1}} + \dots + a_1 \frac{dy(\tau)}{d\tau} \right), \quad (2)$$

$$B(\tau) = \left( b_m \frac{d^m x(\tau)}{d\tau^m} + b_{m-1} \frac{d^{m-1} x(\tau)}{d\tau^{m-1}} + \dots + b_1 \frac{dx(\tau)}{d\tau} \right). \quad (3)$$

Тогда уравнение модели (1) можно записать в виде

$$b_0 \cdot x(\tau) - y(\tau) = A(\tau) - B(\tau). \quad (4)$$

Для получения описания объекта с множеством входных сигналов допустим, что влияние каждого из возмущений на выходную переменную проявляется независимо. Тогда, с учетом

принципа суперпозиции, производные различных порядков выходной переменной по времени определяются как суммы результатов действия различных возмущений. Это позволяет представить модель объекта в виде системы уравнений:

$$\begin{cases} b_{0j} \cdot x_j(\tau) - y(\tau) = B_j(\tau) - A_j(\tau), \\ A_j(\tau) = \left( a_{nj} \left[ \frac{d^n y(\tau)}{d\tau^n} \right]_j + a_{(n-1)j} \left[ \frac{d^{n-1} y(\tau)}{d\tau^{n-1}} \right]_j + \dots + a_{1j} \left[ \frac{dy(\tau)}{d\tau} \right]_j \right), \\ B_j(\tau) = \left( b_{mj} \left[ \frac{d^m x(\tau)}{d\tau^m} \right]_j + b_{(m-1)j} \left[ \frac{d^{m-1} x(\tau)}{d\tau^{m-1}} \right]_j + \dots + b_{1j} \left[ \frac{dx(\tau)}{d\tau} \right]_j \right), \\ \frac{dy^{l \neq 0}(\tau)}{d\tau^{l \neq 0}} = \sum_{i=1}^k \left[ \frac{dy^{l \neq 0}(\tau)}{d\tau^{l \neq 0}} \right]_i, \end{cases} \quad (5)$$

где  $k$  — число возмущений.

Настройка модели (5) по данным о динамике изменения входных сигналов  $X = \{x_1, x_2, \dots, x_n\}$  и выходного сигнала  $y$  во времени производится путем подбора коэффициентов  $a_j$  и  $b_j$ .

Простейшим вариантом модели при наличии единственного входного сигнала является апериодическое звено первого порядка, что предполагает представление модели объекта в виде

$$b_0 \cdot x(\tau) - y(\tau) = a_1 \frac{dy(\tau)}{d\tau}, \quad (6)$$

где  $a_1, b_0$  — параметры настройки, которые должны быть определены с использованием ретроспективных данных.

## 9.2. Проблема настройки модели, оперирующей фактическими значениями сигналов

Сложность настройки модели заключается в неизвестном дрейфе контролируемых значений  $x(\tau)$  и  $y(\tau)$ . Вследствие этого для каждого отдельного периода времени будут получены свои оптимальные значения параметров модели  $a_j$  и  $b_j$ . При этом значения параметров, найденные по данным одного периода времени, могут привести к значительной ошибке в другие периоды.

Разработка глобальной модели требует согласования данных. При этом необходимо обеспечить согласование данных как для каждого периода времени в отдельности, так и данных разных временных периодов.

Решение задачи согласования может быть получено в рамках поисковой процедуры, включающей два этапа. На первом этапе определяем множество возможных решений индивидуально для каждого отдельного периода времени. На втором этапе находим пересечение множеств возможных решений для получения глобальной модели.

Если задача согласования данных в рамках отдельного сегмента времени имеет не единственное решение, то число разных возможных вариантов решений следует ограничивать для последующего согласования данных различных периодов времени при минимальной вычислительной нагрузке.

Если модель оперирует  $x(\tau), y(\tau)$ , то действие неизвестных возмущений в рамках отдельного периода времени может быть представлено некоторой динамикой изменения набора сигналов во времени. Поиск подобной динамики, особенно при наличии нескольких принятых сигналов, описывающих неизвестные воздействия, с высокой вероятностью приведет к неоднозначности решения. Более того, возможно получение значительного числа абсолютно разных решений. В такой ситуации поиск решения, которое позволило бы согласовать данные разных временных периодов, затруднено высокими требованиями к вычислительной нагрузке.

В случае если модель оперирует абсолютными значениями сигналов, то действие низкочастотных составляющих неизвестных возмущений проявляется в смещении производных разных порядков по времени от выходной переменной  $y$ , а также самой выходной переменной  $y$ :

$$\frac{dy^{l \neq 0}(\tau)}{d\tau^{l \neq 0}} = \sum_{i=1}^n \left[ \frac{dy^{l \neq 0}(\tau)}{d\tau^{l \neq 0}} \right]_i + \delta_l, \quad (7)$$

$$b_{0j} \cdot x_j(\tau) - [y(\tau) + \delta_0] = A_j(\tau) - B_j(\tau). \quad (8)$$

В выражениях (7) и (8) коэффициенты  $\delta$  определяют смещение прогноза, связанное с действием неизвестных возмущений. Задача настройки глобальной модели заключается в определении коэффициентов  $a$ ,  $b$ ,  $\delta$  с использованием данных, отобранных за множество отдельных периодов времени. Причем значения коэффициентов  $a$ ,  $b$  должны быть одинаковыми для всех интервалов времени, а значения коэффициентов  $\delta$  — индивидуальными для каждого интервала.

Для снижения вычислительной нагрузки число коэффициентов  $\delta$  может быть снижено путем перехода к приращениям значений сигналов.

### 9.3. Представление сигналов модели в виде приращений относительно начального для периода момента времени

Допустим, что в начальный момент некоторого периода времени объект, обладающий свойством самовывравнивания, находится в состоянии равновесия. Возникновение состояния равновесия можно определить по стабильности значений выходного сигнала  $y$  и входных сигналов  $x$ . Тогда для начального момента времени можем записать

$$A_j(0) = 0, \quad B_j(0) = 0, \quad (9)$$

$$b_{0j} \cdot x_j(0) - y(0) = 0. \quad (10)$$

Учитывая, что, согласно (5),

$$b_{0j} \cdot x_j(\tau) - y(\tau) = B_j(\tau) - A_j(\tau),$$

можем записать

$$B_j(\tau) - A_j(\tau) = b_{0j}(x_j(\tau) - x_j(0)) - (y(\tau) - y(0)). \quad (11)$$

Обозначим величины  $x_j(\tau) - x_j(0)$  и  $y(\tau) - y(0)$  как  $\Delta x_j(\tau)$  и  $\Delta y(\tau)$ . Тогда (11) можно записать в виде

$$B_j(\tau) - A_j(\tau) = k_j \Delta x_j(\tau) - \Delta y(\tau). \quad (12)$$

Если в качестве модели используется апериодическое звено первого порядка, то модель примет вид

$$b_0 \cdot \Delta x(\tau) - \Delta y(\tau) = a_1 \frac{dy(\tau)}{d\tau}. \quad (13)$$

Переход к выражению (13) позволяет исключить из модели коэффициент  $\delta_0$ , который характеризует неопределенное для сегмента данных смещение самой выходной переменной  $y$ , поскольку эффект от приращения значения входа полагаем независимым от уровня выходного сигнала.

Наличие прочих коэффициентов  $\delta$  позволяет учесть нарушение гипотезы об установившемся режиме в начальный момент времени. При гипотезе о наличии у неизвестных возмущений только трендовых составляющих все  $\delta_l$  при  $l > 1$  могут быть приняты равными нулю. Тогда, если в качестве модели используется апериодическое звено первого порядка, модель примет вид

$$a_1 \frac{dy(\tau)}{d\tau} = b_0 k \cdot \Delta x(\tau) - \Delta y(\tau) + \delta_1. \quad (14)$$

#### 9.4. Структура метода

Предлагаемый метод включает ряд этапов, предполагающих определенное преобразование накопленной ретроспективной информации и настройку параметров глобальной модели. Общая структура метода показана на рис. 2.

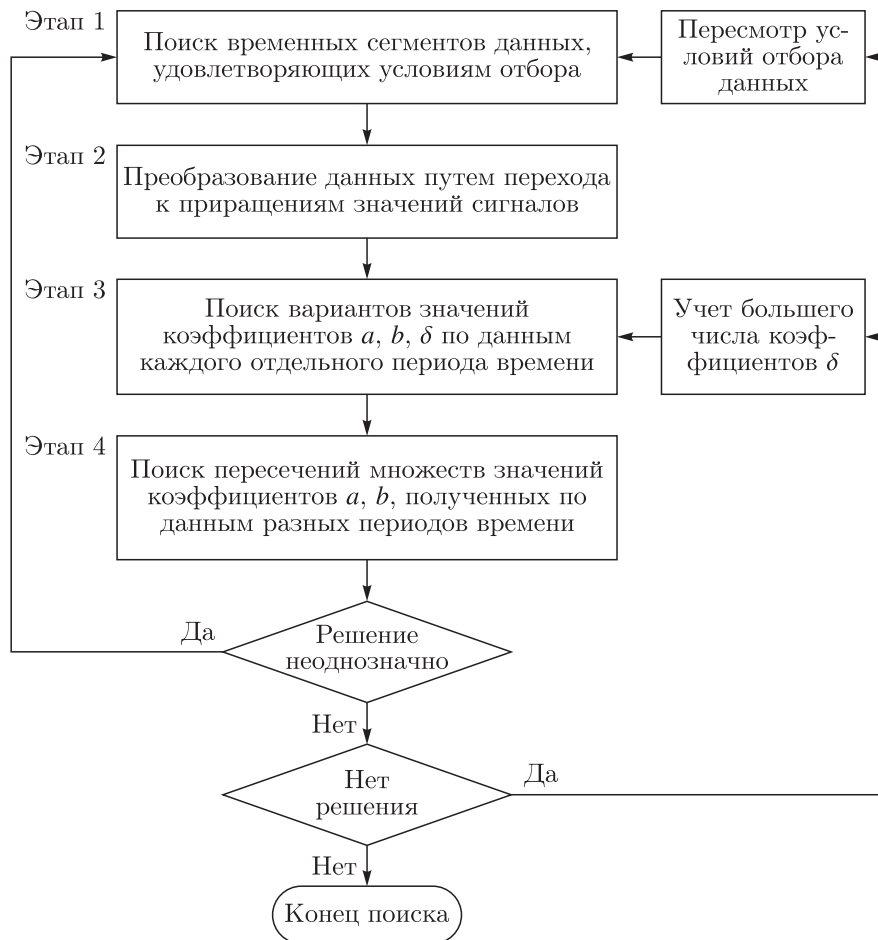


Рис. 2. Структура метода идентификации глобальной модели объекта с самовыравниванием по данным, накопленным за значительный период времени в условиях неизвестных возмущений

На первом этапе отбираются данные периодов времени, для которых выполняются два условия. Первое условие — наличие больших по амплитуде входных воздействий (изменений входных переменных  $x$ ). Второе условие — по возможности стабильность процесса перед воздействиями. Первое условие направлено на снижение влияния на динамику выхода компонентов неизвестных возмущений, которые отличаются от низкочастотных. Поскольку на практике сложно гарантировать наличие у неизвестных возмущений только низкочастотных составляющих, то выбираем периоды времени в данных, для которых величины  $\Delta x(\tau)$  велики. Это позволяет допустить минимальное влияние на динамику выхода компонентов возмущений, отличных от низкочастотных. Второе условие направлено на отбор данных периодов времени, где неизвестные возмущения представлены преимущественно низкочастотными составляющими. Если за некоторое время перед известными воздействиями состояние процесса близко к стабильному при условии отсутствия известных воздействий, то неизвестные возмущения имеют преимущественно низкочастотные компоненты.

Временные пределы периодов целесообразно принять с учетом частотных особенностей известных возмущений и гипотез относительно пределов возможных значений искомых параметров настройки глобальной модели.

Следующий этап — преобразование данных каждого отдельного отобранного периода времени путем перехода к приращениям по отношению к значениям сигналов в начальный для этапа момент времени.

Далее производится поиск значений параметров настройки глобальной модели. Поиск выполняется в два этапа.

Первоначально модель настраивается индивидуально по данным каждого отдельного периода времени. При этом определяется множество возможных значений коэффициентов  $a$  и  $b$  в зависимости от значений  $\delta$ , характеризующих действие неизвестных возмущений. Задача существенно усложняется при большом числе коэффициентов  $\delta$ , и поиск решения целесообразно начинать при единственном для интервала коэффициенте, характеризующем влияние неопределенности на скорость изменения выходной переменной ( $l = 1$ ).

Далее рассматривается пересечение множеств возможных значений  $a$  и  $b$  для разных периодов времени. Данные новых периодов времени подключаются к настройке модели до тех пор, пока сохраняется неопределенность в значениях  $a$  и  $b$  (пересечение множеств дает широкий разброс значений  $a$  и  $b$ ). Если множество значений  $a$  и  $b$  пустое, то возможно два вида деятельности.

Если вычислительные мощности допускают рост вычислительных затрат, то первоначально следует увеличить число коэффициентов  $\delta$ , характеризующих неопределенность. Далее следует повторить оба этапа поиска значений параметров настройки модели. При отсутствии результата от этого мероприятия следует пересмотреть условия отбора данных по двум указанным ранее условиям и повторно сформировать исходные данные для настройки модели.

Отсутствие решения может быть также обусловлено нелинейным влиянием входных переменных  $x$  на выходную переменную  $y$ .

### **9.5. Особенности применения метода при нелинейной модели**

При наличии нелинейных связей между входными переменными  $x$  и выходной переменной  $y$  эффект от приращения значения входного сигнала может быть различен при разных значениях входного или выходного сигнала. Это затрудняет настройку моделей, оперирующих приращениями сигналов.

Следует отметить, что предложенная последовательность этапов не накладывает каких-либо ограничений на вид зависимостей, используемых в модели. Вместо модели (13) может использоваться зависимость любого вида, оперирующая приращениями сигналов. При этом вид зависимостей должен обеспечивать ограниченное число решений по данным отдельного периода времени для работоспособности метода. Для этого модель должна быть интерпретируемой. В качестве модели, например, сложно использовать искусственную нейронную сеть из-за чрезмерно большого числа возможных решений.

Если последовательность этапов на рис. 2 приводит к отсутствию решения предположительно в силу нелинейности, то возможно три мероприятия. Первым мероприятием является расширение множества отобранных решений  $\{a, b, \delta\}$  в рамках третьего этапа путем уменьшения требований к точности представления моделью динамики  $y$  в рамках каждого отобранного интервала времени. Тем самым будет получена модель, которая усредняет влияние приращений значений входных сигналов  $x$  на выход  $y$ . При отсутствии существенных нелинейностей такая модель может быть приемлема. Вторым мероприятием является переход к совокупности моделей, каждая из которых настраивается для набора определенных состояний объекта.



При наличии нелинейных связей между входными переменными  $x$  и выходной переменной  $y$  объекта нарушается принятое допущение о суперпозиции применительно к действию возмущений на выходную переменную. Поэтому третьим мероприятием является повышение требований к стабильности процесса перед известными возмущениями при отборе данных в рамках первого этапа. В противном случае в соответствии с [Pannocchia, Rawlings, 2003] для получения решения может потребоваться иное структурное представление в модели действия неизвестных возмущений.

## 10. Пример идентификации модели охлаждения стали при оцинковании полосы

При термической обработке стали на агрегатах непрерывного горячего оцинкования получили распространение режимы, при которых требуемая температура стальной полосы обеспечивается при перерасходе энергии и при сниженной производительности. Одной из причин этого являются сложности разработки модели, которая могла бы использоваться при упреждающем управлении.

Управление температурой полосы на выходе отделения закрытого охлаждения сопряжено с дополнительной сложностью, связанной с отсутствием контроля температуры рабочего пространства в этом отделении. Для определения рациональной на текущий момент производительности в условиях неизвестных возмущений модель охлаждения полосы должна обеспечивать прогноз температуры полосы на выходе из отделения  $T_M$  с учетом: температуры полосы на входе в отделение  $T_{M0}$ , мощности вентиляторов  $P$ , скорости движения  $v$ , толщины  $h$  и ширины полосы  $b$ . Закономерности модели могут быть определены с использованием теплового баланса отделения за период времени  $\Delta\tau$ :

$$\Delta\tau(Px_1 + T_{\Pi}x_2) = zc_M\Delta\tau(T_M - T_{M0}) + \Delta T_{\Pi}x_3, \quad (15)$$

где  $z = hvb$ ,  $c_M$  — удельная тепловая емкость стали,  $x_1$ ,  $x_2$ ,  $x_3$  — параметры настройки,  $T_{\Pi}$  — неизвестная температура рабочего пространства в отделении,  $T_{\Pi}x_1$  — тепловые потери отделения, не связанные с работой системы охлаждения (приняты пропорциональными  $T_{\Pi}$ ).

Из (15) следует, что сумма тепла  $\Delta\tau Px_1$ , отнятого системой охлаждения, и тепловых потерь  $x_1\Delta\tau T_{\Pi}$  отделения равна сумме тепла, отданного металлом при охлаждении, и тепла, связанного с изменением температурного состояния отделения  $x_3\Delta T_{\Pi}$ .

Выражение (15) содержит неизвестную температуру рабочего пространства  $T_{\Pi}$ . В то же время известной является температура металла на входе и выходе отделения. Для получения оценки  $T_{\Pi}$  в [Ryabchikov et al., 2021] было предложено определить связь между температурой рабочего пространства и температурой металла в виде уравнения теплообмена в конвективной форме:

$$\frac{dT_M(\tau)}{d\tau} = a(T_{\Pi} - T_M(\tau)), \quad a = \frac{\alpha}{h}, \quad (16)$$

где  $\alpha$  — настраиваемый параметр.

Было принято, что в течение времени прохождения сегмента полосы через отделение температура рабочего пространства  $T_{\Pi}$  постоянна. Тогда решение (16) относительно  $T_{\Pi}$  примет вид

$$T_{\Pi} = \frac{T_M(\tau_{\text{охл}}) - T_M(0) + a\tau_{\text{охл}}T_M(\tau_{\text{охл}})}{a\tau_{\text{охл}}}, \quad \tau_{\text{охл}} = \frac{L}{v_1}, \quad (17)$$

где  $T_M(0)$ ,  $T_M(\tau_{\text{охл}})$  — температура полосы на входе и выходе отделения,  $L$  — длина полосы в отделении закрытого охлаждения.

Особенности настройки (17) рассмотрены в работе [Ryabchikov et al., 2021]. Модель может использоваться для прогноза температуры полосы при текущей мощности вентиляторов по температуре металла на входе и выходе отделения, но не позволяет оценить возможность изменения производительности, поскольку не учитывает мощность вентиляторов. В то же время оценка изменения  $T_{\text{п}}$ , полученная с использованием (17), может использоваться при настройке (15). Для этого, используя (17), представим модель (15) в виде

$$\Delta\tau (Px_1^* + T_{\text{п}}x_2^*) = C\Delta\tau(T_{\text{п}} - t_{\text{м0}}) + \Delta T_{\text{п}}x_3^*, \quad (18)$$

где  $C = \frac{bLhv}{aL+hv}$ ,  $x_1^* = \frac{x_1}{ac_{\text{м}}}$ ,  $x_2^* = \frac{x_2}{ac_{\text{м}}}$ ,  $x_3^* = \frac{x_3}{ac_{\text{м}}}$ .

Поделив обе части (18) на  $\Delta\tau$  при  $\Delta\tau \rightarrow 0$ , запишем

$$P(\tau)x_1^* + T_{\text{п}}(\tau)x_2^* = C(T_{\text{п}}(\tau) - t_{\text{м0}}(\tau)) + x_3^* \frac{dT_{\text{п}}(\tau)}{d\tau}. \quad (19)$$

Для настройки модели (19) были отобраны данные периодов, в течение которых наблюдались значимые изменения мощности вентиляторов  $P$  или температуры полосы на входе в отделение  $t_{\text{м0}}$ . Критерием настройки являлась минимизация отклонения прогноза изменения  $T_{\text{п}}$  от оценки  $T_{\text{п}}$ , полученной с применением настроенной модели (17).

Однако попытки настройки модели (19) показали, что соответствие между ретроспективными данными и прогнозом от модели достигается только при индивидуальных настройках модели для каждого периода времени. Определение компромиссных настроек привело к значительной ошибке прогноза и нецелесообразности управления с прогнозирующей моделью.

Применим разработанный метод настройки.

Допустим, что в начальный момент периода температура рабочего пространства в отделении охлаждения установилась. Тогда можем записать (19) для  $\tau = 0$ :

$$P(0)x_1^* + T_{\text{п}}(0)x_2^* = C(T_{\text{п}}(0) - t_{\text{м0}}(0)). \quad (20)$$

Вычитая из (19) выражение (20), получим

$$\Delta P(\tau)x_1^* + \Delta T_{\text{п}}(\tau)x_2^* = C(\Delta T_{\text{п}}(\tau) - \Delta t_{\text{м0}}(\tau)) + x_3^* \frac{dT_{\text{п}}(\tau)}{d\tau}, \quad (21)$$

где  $\Delta P(\tau) = P(\tau) - P(0)$ ,  $\Delta T_{\text{п}}(\tau) = T_{\text{п}}(\tau) - T_{\text{п}}(0)$ ,  $\Delta t_{\text{м0}}(\tau) = t_{\text{м0}}(\tau) - t_{\text{м0}}(0)$ .

Исходя из (21), для скорости изменения температуры рабочего пространства запишем

$$\frac{dT_{\text{п}}(\tau)}{d\tau} = \frac{x_2^* - C}{x_3^*} \left( \frac{x_1^*}{x_2^* - C} \Delta P(\tau) - \Delta T_{\text{п}}(\tau) \right) + \frac{C}{x_3^*} \Delta t_{\text{м0}}(\tau) + x_4^*, \quad (22)$$

где  $x_4^*$  настраивается для учета ошибочности допущения об установившейся температуре в отделении при  $\tau = 0$ .

Настройку модели (22) осуществляли полным перебором значений  $x_1^*$ ,  $x_2^*$ ,  $x_3^*$ ,  $x_4^*$ . Найдено однозначное решение  $\{x_1^* = -35,2, x_2^* = -40,45, x_3^* = 3988,1\}$ , которое позволяет обеспечить наибольшую точность моделирования при выборе для каждого отобранного периода работы агрегата индивидуального значения  $x_4^*$ . При расчетах мощность задавали в процентах от максимума, а значения прочих переменных принимали в формате СИ. На рис. 3 показан пример результатов настройки (22) при тестировании по данным одного из периодов времени.

Модели (17) и (22) могут использоваться совместно при прогнозирующем управлении с варьированием производительности.

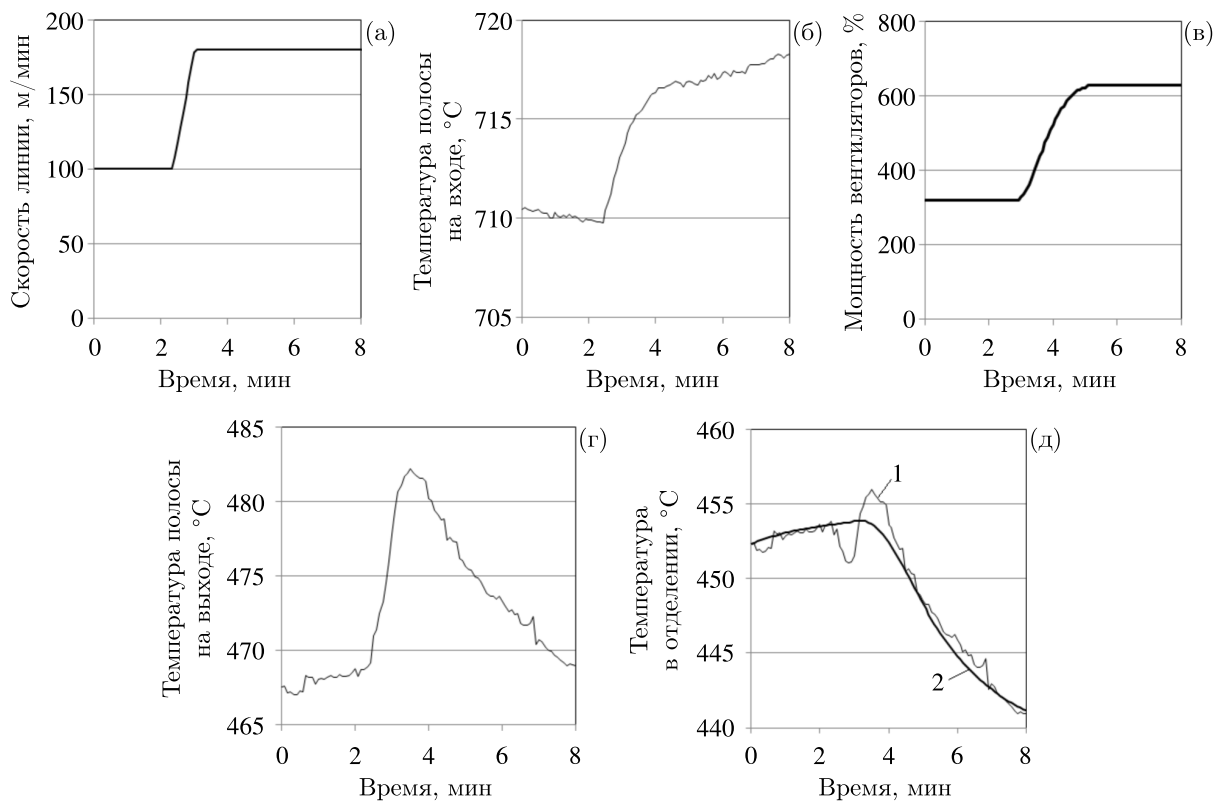


Рис. 3. Результаты моделирования по данным одного из отобранных периодов времени при  $x_4^* = 0,015$ : 1, 2 — модели (17) и (22) соответственно

### 10.1. Особенности организации настройки модели

В рамках первого этапа предложенного метода были отобраны ретроспективные данные шестнадцати периодов времени.

По каждому из четырех параметров  $x_1^*$ ,  $x_2^*$ ,  $x_3^*$ ,  $x_4^*$  приняты двадцать пять уровней варьирования. Далее был проведен автоматизированный вычислительный эксперимент, в ходе которого были получены оценки отклонения расчетов по модели от ретроспективных данных. Всего для каждого отобранного периода были рассчитаны интегральные оценки отклонения (средняя абсолютная ошибка  $E$ ) для  $25^4 = 390\,625$  вариантов значений параметров  $x_1^*$ ,  $x_2^*$ ,  $x_3^*$ ,  $x_4^*$ .

Далее были определены минимумы  $E$  для каждого интервала времени  $\{E1_{\min}, \dots, E16_{\min}\}$  при индивидуальных для интервала значениях  $x_1^*$ ,  $x_2^*$ ,  $x_3^*$ ,  $x_4^*$ . Для изучения влияния допустимого уровня ошибки на поиск настроек глобальной модели был принят мультипликативный коэффициент допустимого повышения ошибки  $k_{\text{er}}$ . Решением считается вариант настроек глобальной модели, при котором для каждого интервала выполняется условие

$$E < k_{\text{er}} \cdot E_{\min}. \quad (23)$$

Если к стационарным параметрам настройки глобальной модели отнести не только  $x_1^*$ ,  $x_2^*$ ,  $x_3^*$ , но и  $x_4^*$ , то решения существуют при  $k_{\text{er}} > 8$ . Такие решения бесполезны в силу низкой точности.

Для определения числа решений в случае, когда к стационарным параметрам настройки глобальной модели относятся только  $x_1^*$ ,  $x_2^*$ ,  $x_3^*$ , а  $x_4^*$  имеет индивидуальное для каждого интервала значение, использовали SQL-оператор вида

```
SELECT count(*)
```

```

FROM (select x1, x2, x3
      from data
      group by x1, x2, x3
      having (min(E1) < Ker*E1_MIN) and
            (min(E2) < Ker*E2_MIN) and
            ...
      )

```

Единственное решение получено при  $k_{er} = 1,5$ . При  $k_{er} = 2$  число решений равно четырём. Однако эти решения находятся в окрестностях решения при  $k_{er} = 1,5$ , и, следовательно, найденное решение можно считать однозначным.

Было определено, что для получения решения при  $k_{er} = 1,5$  достаточно данных шести из шестнадцати отобранных периодов времени. Для прочих десяти периодов решение удовлетворяет условиям вида (23). В таблице 1 показан вариант уменьшения числа решений при последовательном (в произвольном порядке) добавлении ограничений вида (23) для шести периодов.

Таблица 1. Влияние ограничений на количество решений при  $k_{er} = 1,5$

Число интервалов с ограничением вида (23)	1	2	3	4	5	6
Число решений	692	86	19	5	2	1

Полученные результаты показали существенную неравноценность данных разных периодов времени для получения решения. При этом оценить полезность тех или иных данных изначально затруднительно, поскольку влияние данных разных интервалов на число решений проявляется совместно. Возможность совместного использования данных различных временных периодов позволила включить в выборку различные ситуации и минимизировать последствия дисбаланса объема данных по разным режимам работы агрегата. Это связано с тем, что данные периодов, которые не несут новой информации, не приводят к уменьшению числа решений, удовлетворяющих ограничениям.

## Заключение

Предложенный метод идентификации модели со стационарными параметрами ориентирован на вовлечение в решение задач управления больших объемов ретроспективной информации, накопленной за значительный период времени. Большие данные являются мощным ресурсом повышения эффективности управления технологическими процессами и производствами. Разносторонняя информация, накопленная об объекте за длительное время, потенциально позволяет создавать сложные, точные и при этом адекватные в статистическом смысле модели. При управлении подобные модели могут использоваться в случае быстрого изменения состояния объекта и при прогнозе оценок эффективности процессов и производства в целом с участием моделируемого объекта. Однако разработка подобных моделей может быть затруднена вследствие неизвестных возмущений с неопределенными частотными свойствами. В такой ситуации задача одновременного определения стационарных параметров модели и параметров временного ряда неизвестных возмущений в течение всего периода времени, который охватывают большие данные, имеет чрезмерно высокую размерность.

Предложенное в рамках разработанного метода совместное применение отбора данных и представление данных отдельных периодов времени в виде приращений относительно начального момента времени позволяет описать действие неизвестных возмущений одним параметром для каждого отобранного интервала времени. Это способствует снижению размерности поисковой задачи и минимизирует вычислительные затраты, связанные с поиском стационарных параметров модели.

Метод может использоваться не только для линейных, но и для нелинейных моделей. При этом будет получена модель, которая усредняет влияние приращений значений входных сигналов на выход. При отсутствии существенных нелинейностей такая модель может быть приемлема.

Перспективы дальнейшего развития метода связаны с введением дополнительного этапа предварительной сравнительной оценки полезности данных для различных отобранных интервалов времени. Это требует разработки специализированных критериев с применением типовых методов кластеризации. Совершенствование системы отбора данных позволит дополнительно снизить вычислительные затраты.

Предложенный метод продемонстрировал эффективность при решении задачи моделирования охлаждения стальной полосы. Однако необходимо учитывать, что пока применение метода ограничено интерпретируемыми моделями с относительно небольшим числом настроечных параметров. Это существенно затрудняет извлечение пользы из больших данных, поскольку одним из основных преимуществ разносторонней информации о работе объекта, по мнению авторов, является возможность дальнейшего усложнения модели. Вследствие этого одной из основных сфер применения разработанного метода может стать настройка интерпретируемых компонентов гибридных моделей.

## Список литературы (References)

- Казаринов Л. С., Хасанов Д. Р.* Интерактивный метод нейроруавления процессом измельчения шихты в цементном производстве // Вестник Воронежского государственного технического университета. — 2019. — Т. 15, № 6. — С. 46–53.  
*Kazarinov L. S., Khasanov D. R.* Interaktivnyj metod nejroupravleniya processom izmel'cheniya shihty v cementnom proizvodstve [Interactive method of neural control of process of milling mixture in cement production] // The Bulletin of Voronezh State Technical University. — 2019. — Vol. 15, No. 6. — P. 46–53 (in Russian).
- Краснова С. А.* Оценивание производных внешних возмущений на основе виртуальных динамических моделей // Управление большими системами: сборник трудов. — 2018. — № 76. — С. 6–25.  
*Krasnova S. A.* Ocenivanie proizvodnyh vneshnih vozmushchenij na osnove virtual'nyh dinamicheskikh modelej [Estimation of external disturbances derivatives based on virtual dynamic models] // Large-scale Systems Control. — 2018. — Vol. 76. — P. 6–25 (in Russian).
- Райбман Н. С.* Идентификация объектов управления (обзор) // Автоматика и телемеханика. — 1979. — № 6. — С. 80–93.  
*Raibman N. S.* Identification of control objects (review) // Autom. Remote Control. — 1979. — Vol. 40, No. 6. — P. 845–856. (Original Russian paper: *Raibman N. S.* Identifikaciya ob'ektov upravleniya (obzor) // Avtomatika i telemekhanika. — 1979. — No. 6. — P. 80–93.)
- Рябчиков М. Ю., Рябчикова Е. С.* Синтез систем управления с прогнозирующей моделью группой специалистов различного профиля // Вестник Воронежского государственного университета. Серия: Системный анализ и информационные технологии. — 2022. — № 4. — С. 52–73.  
*Ryabchikov M. Yu., Ryabchikova E. S.* Sintez sistem upravleniya s prognoziryushchej model'yu gruppoy specialistov razlichnogo profilya [Synthesis of control systems with a predictive model by a group of specialists of different profiles] // Proceedings of Voronezh State University. Series: Systems Analysis and Information Technologies. — 2022. — No. 4. — P. 52–73. — DOI: 10.17308/sait/1995-5499/2022/4/52-73 (in Russian).
- Рябчиков М. Ю., Рябчикова Е. С., Сунаргулова А. И.* Сэмплинг технологической информации при управлении металлургическими процессами // Автоматизированные технологии и производство. — 2016. — Т. 12, № 2. — С. 34–40.  
*Ryabchikov M. Yu., Ryabchikova E. S., Sunargulova A. I.* Sempling tekhnologicheskoy informacii pri upravlenii metallurgicheskimi processami [Sampling in technical data acquisition tasks] // Automated technologies and production. — 2016. — Vol. 12, No. 2. — P. 34–40 (in Russian).
- Andrievsky B. R., Furtat I. B.* Disturbance observers: methods and applications. I. Methods // Automation and Remote Control. — 2020. — Vol. 81, No. 9. — P. 1563–1610.
- Aved'yan E. D., Tsypkin Y. Z.* Kalman estimation with Brownian disturbances // J. Franklin Inst. — 1998. — Vol. 335B, No. 7. — P. 1195–1205.

- Bandara K., Bergmeir C., Hewamalage H.* LSTM-MSNet: leveraging forecasts on sets of related time series with multiple seasonal patterns // *IEEE Transactions on Neural Networks and Learning Systems*. — 2021. — Vol. 32, No. 4. — P. 1586–1599. — DOI: 10.1109/TNNLS.2020.2985720
- Drovandi C. C., Holmes C. C., McGree J. M., Mengersen K., Richardson S., Ryan E. G.* Principles of experimental design for big data analysis // *Statist. Sci.* — 2017. — Vol. 32, No. 3. — P. 385–404. — DOI: <https://doi.org/10.1214/16-STS604>
- Elragal A., Klischewski R.* Theory-driven or process-driven prediction? Epistemological challenges of big data analytics // *J. Big Data*. — 2017. — Vol. 4, No. 19. — DOI: <https://doi.org/10.1186/s40537-017-0079-2>
- Feng D., Deng Z., Wang T., Liu Y., Xu L.* Identification of disturbance sources based on random forest model // *2018 International Conference on Power System Technology*. — 2018. — P. 3370–3375.
- Francis B. A., Wonham W. M.* The internal model principle of control theory // *Automatica*. — 1976. — Vol. 12, Iss. 5. — P. 457–465.
- Hajiha M., Liu X., Hong Y.* Degradation under dynamic operating conditions: Modeling, competing processes and applications // *Journal of Quality Technology*. — 2021. — Vol. 53, No. 4. — P. 347–368. — DOI: 10.1080/00224065.2020.1757390
- Kazarinov L., Khasanov D.* Interactive mill control // *Proceedings 2019 IEEE Russian Workshop on Power Engineering and Automation of Metallurgy Industry: Research and Practice (PEAMI 2019)*. — 2019. — P. 8–12.
- Kozjek D., Vrabič R., Kralj D., Butala P., Lavrač N.* Data mining for fault diagnostics: A case for plastic injection molding // *Procedia CIRP*. — 2019. — Vol. 81. — P. 809–814.
- Lemeshko B. Yu., Lemeshko S. B., Semenova M. A.* To question of the statistical analysis of big data // *Tomsk State University Journal of Control and Computer Science*. — 2018. — Vol. 44. — P. 40–49.
- Linkan B., Nagi G.* Stochastic modeling and real-time prognostics for multi-component systems with degradation rate interactions // *IIE Transactions*. — 2014. — Vol. 46, No. 5. — P. 470–482. — DOI: 10.1080/0740817X.2013.812269
- Linkan B., Nagi G., Kharoufeh J. P.* Degradation modeling for real-time estimation of residual lifetimes in dynamic environments // *IIE Transactions*. — 2015. — Vol. 47, No. 5. — P. 471–486. — DOI: 10.1080/0740817X.2014.955153
- Malikov A. I.* Synthesis of state unknown inputs observers for nonlinear lipschitz systems with uncertain disturbances // *Automation and Remote Control*. — 2018. — Vol. 79, No. 3. — P. 406–424.
- Maxwell K., Rajabi M., Esterle J.* Automated classification of metamorphosed coal from geophysical log data using supervised machine learning techniques // *International Journal of Coal Geology*. — 2019. — Vol. 214. — Article 103284. — DOI: 10.1016/j.coal.2019.103284
- Pannocchia G., Rawlings J. B.* Disturbance models for offset-free model-predictive control // *AIChE J.* — 2003. — Vol. 49. — P. 426–437. — DOI: <https://doi.org/10.1002/aic.690490213>
- Ryabchikov M. Yu.* Selection of steel strip annealing energy-saving conditions in view of the substandard products share // *Journal of Chemical Technology and Metallurgy*. — 2020. — Vol. 55, No. 1. — P. 182–191.
- Ryabchikov M. Yu., Ryabchikova E. S.* Big data-driven assessment of proposals to improve enterprise flexibility through control options untested in practice // *Glob. J. Flex. Syst. Manag.* — 2022. — Vol. 23. — P. 43–74. — DOI: <https://doi.org/10.1007/s40171-021-00287-5>
- Ryabchikov M. Yu., Ryabchikova E. S., Shmanev D. E., Kokorin I. D.* Strip cooling control for flexible production of galvanized flat steel // *Steel in Translation*. — 2021. — Vol. 51, No. 7. — P. 446–455.
- Shardt Y. A. W., Brooks K.* Automated system identification in mineral processing industries: A case study using the zinc flotation cell // *IFAC-PapersOnLine*. — 2018. — Vol. 51, No. 18. — P. 132–137.
- Wahlberg B.* Laguerre and Kautz models // *IFAC Proceedings Volumes*. — 1994. — Vol. 27, Iss. 8. — P. 965–976.

- 
- Wahlberg B.* System identification using Laguerre models // IEEE transactions on automatic control. — 1991. — Vol. 36, Iss. 5. — P. 551–562.
- Wang C., Zhu Z., Lu N., Cheng Y., Jiang B.* A data-driven degradation prognostic strategy for aero-engine under various operational conditions // Neurocomputing. — 2021. — Vol. 462. — P. 195–207. — DOI: 10.1016/j.neucom.2021.07.080
- Wang H., Liao H., Ma X.* Remaining useful life prediction considering joint dependency of degradation rate and variation on time-varying operating conditions // IEEE Transactions on Reliability. — 2021. — Vol. 70, No. 2. — P. 761–774. — DOI: 10.1109/TR.2020.3002262