

На правах рукописи



Щербаков Артем Петрович

**РАЗРАБОТКА МЕТОДОВ И АЛГОРИТМОВ РЕКУРРЕНТНОЙ
ИДЕНТИФИКАЦИИ ИЕРАРХИЧЕСКИХ ОКРЕСТНОСТНЫХ
МОДЕЛЕЙ**

Специальность 1.2.2 – Математическое моделирование, численные методы и
комплексы программ

АВТОРЕФЕРАТ
диссертации на соискание учёной степени
кандидата технических наук

Липецк – 2024

Работа выполнена в ФГБОУ ВО «Липецкий государственный технический университет»

Научный руководитель: **Шмырин Анатолий Михайлович**,
доктор технических наук, профессор, ФГБОУ
ВО «Липецкий государственный технический
университет», заведующий кафедрой высшей
математики, г. Липецк

Официальные оппоненты: **Матвеев Михаил Григорьевич**,
доктор технических наук, профессор, ФГБОУ
ВО «Воронежский государственный
университет», заведующий кафедрой
информационных технологий управления,
г. Воронеж

Алексеев Владимир Витальевич,
доктор технических наук, профессор, ФГБОУ
ВО «Тамбовский государственный технический
университет», профессор кафедры
информационных систем и защиты информации,
г. Тамбов

Ведущая организация: Федеральное государственное автономное
образовательное учреждение высшего
образования «**Белгородский государственный
национальный исследовательский
университет**», г. Белгород

Защита диссертации состоится 26 марта 2024 года в 14:00 на заседании
диссертационного совета 24.2.323.01 в ФГБОУ ВО «Липецкий
государственный технический университет» по адресу: 398055, г. Липецк, ул.
Московская, 30.

С диссертацией можно ознакомиться в библиотеке ФГБОУ ВО «Липецкий
государственный технический университет» и на сайте www.stu.lipetsk.ru.

Автореферат разослан «07» февраля 2024 г.

Учёный секретарь
диссертационного совета 24.2.323.01
д.т.н., доцент

И.А. Седых

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования и степень её разработанности. Многие современные сложные производственные процессы характеризуются распределенностью во времени и пространстве, нелинейностью и наличием спектра номинальных технологических режимов как производства в целом, так и отдельных его составляющих или узлов. Проявления нелинейности многообразны и, как правило, связаны со спецификой изучаемого процесса или объекта, и потому не существует универсальных методов построения нелинейных моделей. Разработка новых математических методов и алгоритмов, адекватно отражающих характер нелинейности отдельных номинальных режимов или всего моделируемого процесса в целом, остается актуальной задачей, несмотря на большое количество уже имеющихся методов и алгоритмов, даже таких мощных как нейронные сети или ансамблевые алгоритмы на основе деревьев регрессии.

Одним из перспективных направлений моделирования сложных распределенных нелинейных процессов является окрестностное моделирование, основанное на математической формализации технологической схемы процесса в виде окрестностной структуры – орграфа с наборами переменных и операторов, ассоциированных, соответственно, с вершинами и дугами орграфа. Построенной окрестностной структуре соответствует окрестностная система, уравнения которой, в зависимости от рассматриваемой задачи, могут быть статическими или динамическими, линейными или нелинейными и так далее. Теория окрестностных систем и связанные с ней задачи идентификации и управления рассматривались в работах С.Л. Блюмина, А.М. Шмырина, Н.Н. Карабутова, И.А. Седых и других авторов.

Нелинейность в окрестностных моделях связана, прежде всего, с выбором типа уравнений или операторов окрестностной системы. Во многих работах рассматривались билинейные и, в более общем случае, полиномиальные окрестностные системы. Ограниченность области применения полиномиальных окрестностных систем связана с их локальностью (как и формулы Тейлора) и, кроме того, с проблемами мультиколлинеарности, возникающей в задаче идентификации в связи с введением искусственных предикторов, соответствующих нелинейным одночленам. В работах И.А. Седых рассматривались нейроокрестностные модели, в которых нелинейность в узлах (вершинах) моделировалась с помощью нейронных сетей как операторов узлов окрестностной системы, а нелинейность процесса в целом была связана с рассмотрением переменных окрестностей в динамическом случае. Для идентификации таких моделей, включающих нейронные сети, требуется большое количество данных. В диссертации А.С. Канюгиной рассматривалась схема объединения нескольких линейных окрестностных моделей, соответствующих номинальным технологическим режимам производства, в общую нелинейную модель с возможностью перехода между номинальными режимами.

Настоящая работа посвящена дальнейшему развитию методов нелинейного моделирования во взаимосвязи с теорией окрестностных систем.

Уже в первых работах по окрестностному моделированию отмечалось, что окрестностные системы могут быть полезны не только в задачах моделирования производственных процессов, но и как средство интерпретации и анализа многих математических методов и алгоритмов, таких, например, как метод сеток, метод конечных элементов, нейронные сети и сети Петри. Продолжая данное «алгоритмическое» направления применения окрестностных систем, представленные в работе методы и алгоритмы рекуррентной иерархической идентификации нелинейных моделей описываются с помощью иерархических окрестностных структур и систем.

Рассматриваются методы уточнения заданной первоначальной модели (например, модели узлового оператора окрестностной системы) посредством рекуррентной идентификации дополнительных слагаемых на основании использования остаточных данных (невязок) предыдущей модели. Корректирующие слагаемые соответствуют вершинам некоторой иерархической окрестностной структуры, орграф такой структуры называется деревом модели. В первом методе, который можно назвать статическим, дерево модели соответствует заранее заданной иерархической кластеризации множества кортежей входных данных, то есть дерево фиксировано и не изменяется. Во втором методе, который можно назвать динамическим, дерево модели соответствует возникающему в процессе идентификации иерархическому разбиению (фрагментации) множества входных кортежей, то есть дерево определяется в процессе идентификации одновременно с разбиением и не задано заранее. Важно отметить, что, в отличие, например, от нейросетевых методов, данные методы не требуют большого количества экспериментальных данных.

Наличие дерева в методах рекуррентной иерархической идентификации придает им сходство с известным алгоритмом CART (Classification and regression trees), разработанным Л. Брейманом, Д. Фридманом и другими, а рекуррентное использование остаточных данных – с ансамблевым методом бустинга на основе CART. В действительности представленные в работе алгоритмы, несмотря на указанное внешнее сходство, существенно отличаются от алгоритма CART и основанных на CART ансамблевых методах. Анализ сходства и различия алгоритмов удобно проводить, используя их представление в виде (определяемых в данной работе) иерархических окрестностных моделей.

Тематика работы связана с научными направлениями ФГБОУ ВО «Липецкий государственный технический университет» «Исследование и разработка методов и алгоритмов прикладной математики для идентификации технологических и сопровождающих процессов» и «Современные сложные системы управления».

Цель работы и задачи исследования. Целью диссертационной работы является разработка методов и алгоритмов рекуррентной идентификации иерархических окрестностных моделей и анализа остаточных данных и применение разработанных алгоритмов в задачах моделирования процессов с нелинейным поведением. В рамках данного исследования были поставлены и решены следующие задачи:

1. Обзор существующих классов окрестностных моделей и разработка моделей с иерархической окрестностной структурой.

2. Разработка алгоритмов и численных методов рекуррентной идентификации иерархических окрестностных моделей на основе использования остаточных данных выходов (невязок) промежуточных моделей.

3. Разработка метода контроля количества уровней в иерархических окрестностных моделях на основе анализа остатков промежуточных моделей.

4. Применение алгоритмов рекуррентной идентификации иерархических окрестностных моделей для исследования и прогнозирования свойств и характеристик процессов с нелинейным поведением.

5. Разработка проблемно-ориентированных модулей программного обеспечения, реализующих алгоритмы рекуррентной идентификации иерархических моделей.

Научная новизна. В диссертационной работе получены следующие результаты, отличающиеся научной новизной:

1. Введен класс окрестностных моделей, отличающийся иерархической структурой с двусторонними связями и возможностью добавления новых узлов, что позволяет расширить возможности применения окрестностных систем в задачах моделирования производственных процессов и задачах описания иерархических алгоритмов.

2. Разработан алгоритм рекуррентной идентификации иерархических окрестностных моделей на основе аппроксимации невязок промежуточных моделей, отличающийся использованием заданной иерархической кластеризации обучающих входных данных и позволяющий повысить точность моделей.

3. Разработан алгоритм и численный метод рекуррентной идентификации иерархических окрестностных моделей на основе трихотомии невязок промежуточных моделей, отличающийся построением в процессе идентификации иерархического разбиения обучающих входных данных и позволяющий улучшать точность моделей и качество прогноза.

4. Предложен метод анализа остаточных данных промежуточных моделей в алгоритмах рекуррентной идентификации, отличающийся введением порядка на множестве обучающих данных и позволяющий контролировать количество уровней иерархии, необходимое для достижения соответствия между точностью и качеством прогноза иерархической модели.

5. Разработан комплекс проблемно-ориентированных программ, реализующих алгоритмы рекуррентной идентификации иерархических моделей, отличающихся наличием модуля вычисления ближайшего обучающего кортежа и позволяющих оценивать длину дерева модели.

Теоретическая и практическая значимость. Теоретическая значимость результатов работы заключается в разработанных алгоритмах рекуррентной идентификации иерархических окрестностных моделей аппроксимации данных.

Практическая значимость результатов работы заключается в построении моделей прогнозирования модульных характеристик клинкера после обжига в

зависимости от химического состава до обжига и модели прогнозирования температурного режима стадии диффузии производства сахара.

Объект исследования. Производственные процессы с существенно нелинейными характеристиками.

Предмет исследования. Моделирование сложных производственных процессов с существенно нелинейными характеристиками с помощью алгоритмов рекуррентной идентификации иерархических моделей.

Методология и методы исследования. В качестве теоретической и методологической основы выступают методы теории математического моделирования, вычислительной математики, теории графов, математической статистики.

Положения, выносимые на защиту:

1. Новый класс окрестностных моделей с иерархической структурой, позволяющий расширить возможности применения окрестностных систем в задачах описания иерархических алгоритмов и моделирования нелинейных процессов.

2. Алгоритм рекуррентной идентификации иерархических окрестностных моделей на основе аппроксимации невязок промежуточных моделей, отличающийся использованием заданной иерархической кластеризации обучающих входных данных.

3. Алгоритм и численный метод рекуррентной идентификации иерархических окрестностных моделей на основе трихотомии невязок промежуточных моделей, отличающийся построением в процессе идентификации иерархического разбиения обучающих входных данных.

4. Метод контроля количества уровней иерархии в моделях, необходимых для достижения соответствия между точностью и качеством прогноза иерархической модели.

5. Комплекс проблемно-ориентированных программ, реализующих алгоритмы рекуррентной идентификации иерархических моделей и анализа остаточных данных.

Тематика работы соответствует следующим пунктам паспорта специальности 1.2.2:

разработка алгоритма иерархической идентификации на основе аппроксимации невязок промежуточных моделей, разработка алгоритма и численного метода иерархической идентификации на основе трихотомии невязок промежуточных моделей, позволяющих повысить качество аппроксимации и точность прогноза соответствует пункту 2 «Разработка, обоснование и тестирование эффективных вычислительных методов с применением современных компьютерных технологий»;

разработка метода проверки адекватности иерархических моделей, отличающегося анализом остаточных данных промежуточных моделей, упорядоченных по возрастанию обучающих выходов, и позволяющий контролировать количество необходимых уровней иерархии в моделях соответствует пункту 5 «Разработка новых математических методов и алгоритмов валидации математических моделей объектов на основе данных натурального эксперимента или на основе анализа математических моделей»;

исследование окрестностных моделей с иерархической структурой прогнозирования свойств клинкера в цементном производстве и температурного режима стадии диффузии производства сахара соответствует пункту 8 «Комплексные исследования научных и технических проблем с применением современной технологии математического моделирования и вычислительного эксперимента».

Степень достоверности результатов работы. Достоверность результатов работы подтверждается проведенными в достаточном объеме исследованиями с применением современных технологий математического моделирования и вычислительного эксперимента. Разработанные методы, процедуры и алгоритмы были применены для исследования реальных объектов

– химического состава сырьевой муки и клинкера при обжиге клинкера в печах;

– температурных показателей процесса диффузии и экстрагирования производства сахара.

Сравнительный анализ полученных результатов показал соответствие производственным данным.

Реализация и внедрение результатов работы. Результаты диссертационной работы рекомендованы к использованию акционерным обществом «Липецкцемент» для анализа и управления технологическими показателями предприятия; рекомендованы к использованию на структурном подразделении «Хмелинецкий сахарный завод» акционерного общества «Агропромышленное объединение «Аврора» для уточнения и совершенствования существующих математических моделей сложных производственных процессов.

Теоретические результаты диссертации используются в учебном процессе ФГБОУ ВО «Липецкий государственный технический университет» при выполнении курсовых работ по дисциплине «Математическое моделирование», а также подготовке выпускных квалификационных работ.

Апробация работы. Теоретические и практические результаты работы докладывались и обсуждались на международных конференциях: Международная научно-техническая конференция «Фундаментальные и прикладные проблемы модернизации современного машиностроения и металлургии» (Липецк, 2012), XVII Международная научная конференция «Современные проблемы информатизации в экономике и обеспечении безопасности» (Воронеж, 2012), Международная научно-практическая конференция «Современные проблемы в образовании и науке» (Тамбов, 2013), VI международная конференция «Современные методы прикладной математики, теории управления и компьютерных технологий» («ПМТУКТ») (Воронеж, 2013), XX-th International Open Science Conference «Modern informatization problems in simulation and social technologies» (Yelm, WA, USA, 2015), 2nd International Conference on Technology Enhanced Learning in Higher Education TELE2022 (Lipetsk, 2022), 4rd International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency SUMMA2022 (Lipetsk, 2022), 5th International Conference on Control Systems, Mathematical

Modeling, Automation and Energy Efficiency SUMMA2023 (Lipetsk, 2023), Международная научно-практическая конференция «Нано-био-технологии. Теплоэнергетика. Математическое моделирование» (Липецк, 2023); всероссийских конференциях: IX всероссийская школа-конференция молодых ученых «Управление большими системами» (Тамбов, 2012), областных региональных конференциях, а также на научных семинарах кафедры высшей математики Липецкого государственного технического университета.

Публикации. Основные результаты работы опубликованы в 26 печатных трудах, в том числе 5 – самостоятельно, из них 5 статей в ведущих рецензируемых журналах, рекомендованных в Перечне ВАК, 2 – в изданиях, входящих в международные системы цитирования Scopus и Web of Science, 3 свидетельства о государственной регистрации программы для ЭВМ.

В работах, написанных в соавторстве и приведенных в автореферате, лично соискателем получены следующие результаты: [2, 12] – разработка и применение на сгенерированных данных алгоритмов рекуррентной идентификации иерархических окрестностных моделей, определение иерархического разбиения и иерархической кластеризации, построение непрерывной модели с помощью иерархического разбиения единицы; [3, 7] – разработка схемы агрегирования локальных мод системы, зависящих от кластеров входов, в общую мультимодальную окрестностную систему; [6] – разработка методов стабилизации локальных мод; [4, 15-17, 20] – определение компонент состояния и управления окрестностной системы печи обжига клинкера, [5, 21, 23] – оценка адекватности общих и параметрических окрестностных систем; [8] – разработка программного обеспечения, реализующего численные методы анализа нелинейности входного процесса; [9] – алгоритмическое обеспечение для получения общего параметрического уравнения окрестностной модели; [10] – алгоритмическое обеспечение для решения систем методом ортонормализации Грама-Шмидта; [19] – анализ нелинейности при исследовании характеристик производства клинкера; [22, 24-26] – разработка окрестностных моделей распределенных динамических систем.

Структура и объем работы. Диссертация состоит из введения, четырёх глав, заключения, списка литературы из 134 наименований и 2 приложений на 6 страницах. Объем основной части работы составляет 134 страницы, включая 24 рисунка и 6 таблиц.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы исследования, сформулированы цели и задачи работы, научная новизна, теоретическая и практическая значимость работы, методология и методы исследования, перечислены положения, выносимые на защиту, степень достоверности и апробации результатов, а также основное содержание работы.

Первая глава содержит обзор различных видов линейных и нелинейных окрестностных моделей, в том числе описанных в литературе классов окрестностных моделей иерархического типа. Указаны два направления

применения окрестностных моделей, технологическое и вычислительное. Представлены сведения об алгоритмах решающих деревьев и основанных на них ансамблевых методах. Обсуждается проблема соотношения точности моделей и качества прогноза.

Вопросами идентификации линейных и нелинейных окрестностных систем занимались С.Л. Блюмин, А.М. Шмырин, Н.Н. Карабутов, И.А. Седых и другие авторы. В работах А.М. Шмырина и И.А. Седых были определены два класса иерархических окрестностных моделей. В первом случае иерархичность модели понималась как многослойность окрестностной структуры с направлением связей от старших слоев к младшим и произвольными связями внутри слоев, то есть каждый отдельно рассматриваемый слой мог быть произвольной окрестностной моделью. Во втором случае иерархичность модели понималась как введение дополнительных окрестностных моделей второго уровня внутри узлов первого уровня.

Окрестностные структуры и системы обычно применяются для построения математических моделей производственных процессов на основе формализации технологических схем этих процессов в виде окрестностных структур. Такое «технологическое» применение можно считать первым направлением в теории окрестностного моделирования.

В рамках второго направления окрестностные модели используются для представления и анализа разнообразных вычислительных методов и алгоритмов. «Вычислительное» направление применения окрестностных систем рассматривалась в работах С.Л. Блюмина, А.М. Шмырина, И.А. Седых и др. применительно к таким методам и алгоритмам, как, например, метод конечных разностей, метод конечных элементов, нейронные сети, сети Петри, и т.д.

Во второй главе определяются иерархические окрестностные модели с окрестностной структурой в виде ориентированного дерева с восходящими и нисходящими дугами или дугами только одного типа. Этот класс иерархических окрестностных моделей не совпадает с ранее определенными классами, описанными в первой главе. Такие модели могут быть использованы в рамках «технологического» направления, например, для описания иерархических организационных структур с передачей отчетов по восходящим связям и циркуляров или приказов по нисходящим. В данной работе эти модели используются для представления разработанных алгоритмов рекуррентной идентификации и для анализа сходства и различия этих алгоритмов с алгоритмом CART и ансамблевым методом бустинга на основе CART. Далее термин «иерархическая окрестностная модель» используется для моделей с древовидной окрестностной структурой.

Описываются иерархические кластеризации и иерархические разбиения обучающих данных, определяются связанные с кластеризациями и разбиениями иерархические окрестностные структуры. Дается интерпретация процесса построения и дальнейшего использования дерева регрессии в алгоритме решающих деревьев (CART) как процесса идентификации и последующего применения иерархической окрестностной модели с нисходящими связями.

Иерархическое разбиение множества кортежей данных $D \subset \mathbb{R}^n$ задается сюръективным отображением $D \rightarrow S$ на множество S листьев (выходов) дерева. Это отображение продолжается до отображения в S ячеек Дирихле-Вороного, соответствующих множеству кортежей данных, и потому порождает иерархическое разбиение пространства \mathbb{R}^n , в котором каждый элемент разбиения любого уровня состоит из некоторого набора ячеек Дирихле-Вороного.

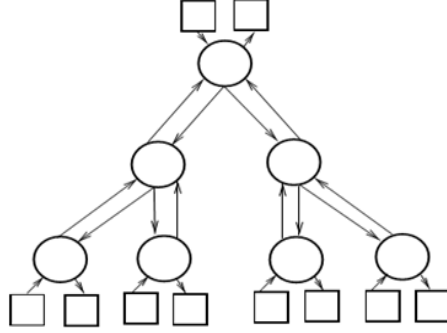


Рисунок 1 – Пример иерархической окрестностной структуры

Для описания иерархической окрестностной модели, соответствующей иерархической окрестностной структуре, вводятся следующие обозначения. Входные и выходные данные корневого узла — это $u(0)$ и $y(0)$. Корневой узел — $x(i_1) = x(1)$. Далее, рекуррентно, $x(i_1, \dots, i_s, i_{s+1})$ является i_{s+1} -м узлом из упорядоченного набора узлов, подчиненных узлу $x(i_1, \dots, i_s)$. Входными и выходными данными терминальных узлов $x(i_1, \dots, i_s, i_{s+1})$ являются $u(i_1, \dots, i_s, 1)$ и $y(i_1, \dots, i_s, 1)$. Количество узлов, подчиненных узлу $x(i_1, i_2, \dots, i_s)$, обозначается $n(i_1, i_2, \dots, i_s)$. Обозначения переменных аналогичны обозначениям узлов; $I_s = (i_1, \dots, i_s)$. Каждая переменная, относящаяся к дуге, направленной вверх (соответственно вниз), обозначается так же, как начальная (соответственно конечная) вершина соответствующей дуги, с заменой буквы на заглавную и индексами u (соответственно d). Оператор, соответствующий узлу $x(i_1, \dots, i_s)$, представляется в виде двух операторов F и G для дуг, направленных вверх и вниз. Уравнения иерархической окрестностной модели с k уровнями имеют вид:

$$\left\{ \begin{array}{l} Y_u(0) = X_u^{t+1}(1) = F_1 \left(X_u^t(1,1), \dots, X_u^t(1, n(1)) \right); \\ \quad [X_d^{t+1}(1,1), \dots, X_d^{t+1}(1, n(1))] = \\ \quad G_1 \left(X_u^t(1,1), \dots, X_u^t(1, n(1)), U_d^t(0) = X_d^t(1) \right); \\ X_u^{t+1}(I_s) = F_{I_s} \left(X_u^t(I_s, 1), \dots, X_u^t(I_s, n(I_s)), X_d^t(I_s) \right); \\ \quad [X_d^{t+1}(I_s, 1), \dots, X_d^{t+1}(I_s, n(I_s))] = \\ \quad G_{I_s} \left(X_u^t(I_s, 1), \dots, X_u^t(I_s, n), X_d^t(I_s) \right); \\ \quad \quad \quad s = 2, \dots, k - 1; \\ X_u^{t+1}(I_k) = F_{I_k} \left(U_u(I_k, 1), X_d^t(I_k) \right); \\ \quad \quad \quad Y_d^{t+1}(I_k, 1) = G_{I_k} \left(X_d^t(I_k) \right). \end{array} \right. \quad (1)$$

Если иерархическая модель содержит только нисходящие (соответственно восходящие) дуги, то уравнения окрестностной модели будут иметь вид:

$$\left\{ \begin{array}{l} [X_d^{t+1}(1,1), \dots, X_d^{t+1}(1, n(1))] = G_1(X_d^t(1)); \\ [X_d^{t+1}(I_s, 1), \dots, X_d^{t+1}(I_s, n(I_s))] = G_{I_s}(X_d^t(I_s)); \\ \quad s = 2, \dots, k - 1; \\ Y_d^{t+1}(I_k, 1) = G_{I_k}(X_d^t(I_k)) \end{array} \right. \quad (2)$$

и, соответственно,

$$\left\{ \begin{array}{l} Y_u(0) = X_u^{t+1}(1) = F_1(X_u^t(1,1), \dots, X_u^t(1, n(1))); \\ X_u^{t+1}(I_s) = F_{I_s}(X_u^t(I_s, 1), \dots, X_u^t(I_s, n(I_s))); \\ \quad s = 2, \dots, k - 1; \\ X_u^{t+1}(I_k) = F_{I_k}(U_u(I_k, 1)), \end{array} \right. \quad (3)$$

В третьей главе описываются статический и динамический алгоритмы рекуррентной идентификации иерархических окрестностных моделей, а также численный метод рекуррентной идентификации, основанный на трихотомии невязок и позволяющий контролировать точность аппроксимации с помощью выбора параметра трихотомии. Предлагается схема контроля момента остановки рекуррентных алгоритмов при достижении соответствия точности и качества прогноза модели на основе анализа остаточных данных выходов уже построенных моделей для данного уровня иерархии. Рассматривается задача синтеза мультимодальных окрестностных систем на основе агрегирования номинальных режимов системы в целом или номинальных режимов отдельных узлов.

Оба алгоритма рекуррентной идентификации иерархических моделей представляют собой схему последовательного уточнения некоторой начальной аналитической модели, при этом окончательный результат интерпретируется как иерархическая окрестностная модель с восходящими дугами. Окрестностная структура иерархической модели (дерево модели) в статическом алгоритме известна заранее и соответствует заданной иерархической кластеризации (или разбиения) множества обучающих данных. В динамическом алгоритме окрестностная структура заранее не известна и определяется в процессе последовательного уточнения модели одновременно с построением иерархического разбиения множества обучающих данных. Элементы разбиения уже построенного множества данного уровня на подмножества следующего уровня определяются как прообразы кластеров соответствующего множества невязок; эти кластеры могут быть заданы, например, с помощью трихотомии невязок по величине и знаку.

Описываемые статический и динамический алгоритмы названы в работе также R-алгоритмом и L-алгоритмом по причине их формального сходства с конструкциями интегралов Римана и Лебега.

Схема R-алгоритма рекуррентной идентификации



Рисунок 2 – Схема R-алгоритма идентификации

1. Задается некоторая иерархическая кластеризация (или, в более общем случае, иерархическое разбиение) множества кортежей входных данных $D \subset \mathbb{R}^n$, кодируемое деревом T , и ассоциированное с ним иерархическое разбиение области определения $U \subset \mathbb{R}^n$ входов модели. Каждый из элементов иерархического разбиения области U состоит из многогранников Дирихле-Вороноя.

2. Идентифицируется некоторая начальная модель $F^1(x)$ по всем данным $D \subset \mathbb{R}^n$.

3. Вычисляются остатки (невязки) начальной модели:

$$\hat{d}^1 = \hat{d} - F^1(d), \quad (4)$$

где $\hat{d} \in \mathbb{R}^1$ – это значение выхода для входа $d \in \mathbb{R}^n$. Идентифицируются уточняющие модели $F_{I_2}^2(x)$ для каждого элемента разбиения второго уровня, $I_2 = (1, i)$ – код вершины дерева второго уровня.

4. Процесс итерируется: для каждой модели, соответствующей элементу разбиения второго уровня D_1, D_2, \dots, D_r , вычисляются остатки и идентифицируются модели для элементов разбиения третьего уровня иерархии и т.д.

Схема L-алгоритма рекуррентной идентификации

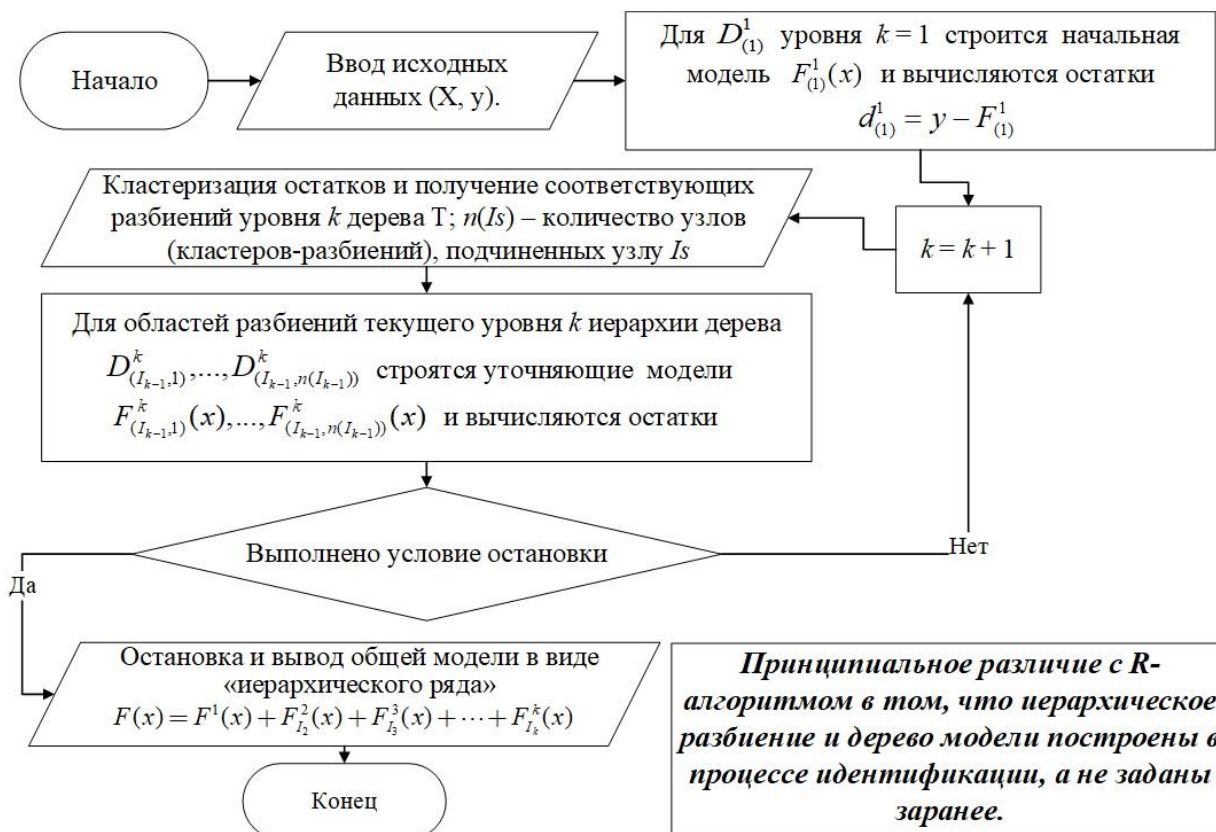


Рисунок 3 – Схема L-алгоритма идентификации

1. Идентифицируется некоторая начальная модель $F^1(x)$ по всем данным.
2. Вычисляются остатки (невязки) начальной модели:

$$\hat{d}^1 = \hat{d} - F^1(d) \quad (5)$$

3. Определяется кластеризация остатков (например, на три кластера с помощью трихотомии). Возникает ассоциированное с кластерами остатков разбиение $D = D_1 \cup \dots \cup D_r$ пространства обучающих входов на прообразы кластеров (в случае трихотомии $r = 3$).

4. Для каждого из подмножеств D_i и соответствующих остатков идентифицируются уточняющие модели второго уровня $F_{I_2}^2(x)$, $I_2 = (1, i)$.

5. Процесс итерируется: для каждой из моделей последнего уровня k (из уже построенных) проводится кластеризация (например, трихотомия) невязок этой модели, находятся прообразы в пространстве обучающих входов для каждого кластера и идентифицируются уточняющие модели следующего уровня.

Итоговая модель для обоих алгоритмов имеет вид «иерархического ряда»:

$$F(x) = F^1(x) + F_{I_2}^2(x) + F_{I_3}^3(x) + \dots + F_{I_k}^k(x) \quad (6)$$

где $I_r = I_r(x)$ – коды вершин, лежащих на пути от концевой вершины, соответствующей кортежу $x \in \mathbb{R}^n$, к корню. Уточняющими моделями $F_{I_r}^r(x)$ на

каждом уровне иерархии могут служить любого вида модели, связывающие входные и выходную величины, например, линейные и полиномиальные регрессионные модели. В случае R-алгоритма с иерархической кластеризацией обучающих входов иерархический ряд можно заменить обычной формулой, используя в качестве весов слагаемых гауссианы соответствующих кластеров. В такой записи R-модель может быть интерпретирована как иерархическое обобщение модели Такаги-Сугено-Канга.

L-идентификация с трихотомией невязок.

Выбирается некоторое число ε (точность аппроксимации) и после построения начальной модели множество входных данных $D \subset \mathbb{R}^n$ разбивается на три подмножества по правилу:

$$\begin{aligned} D &= D_\varepsilon \cup D^- \cup D^+, \\ D_\varepsilon &= \{d \in D; |\hat{d} - F^1(d)| < \varepsilon\}, \\ D^- &= \{d \in D; F^1(d) \geq \hat{d} + \varepsilon\}, \\ D^+ &= \{d \in D; F^1(d) \leq \hat{d} - \varepsilon\}. \end{aligned} \quad (7)$$

Для множества D_ε процесс построения модели закончен (требуемая точность уже достигнута), для множеств D^- и D^+ идентифицируются уточняющие модели.

На рис. 4 представлено схематичное изображение процесса L-идентификации на основе трихотомии невязок. Прообразы кластеров в пространстве входов \mathbb{R}^n (на рисунке – \mathbb{R}^2) изображены как точки одного цвета с соответствующими кластерами.

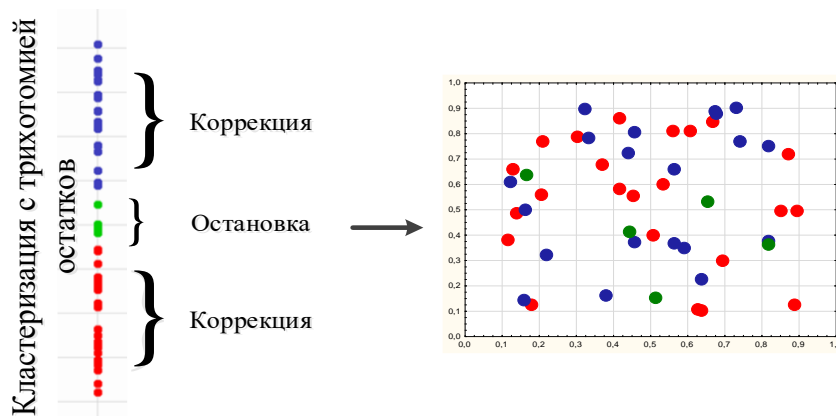


Рисунок 4 – Схематичное изображение L-алгоритма с трихотомией

Идентификация модели при работе обоих алгоритмов происходит «сверху вниз», т.е. от корневой вершины к концевым (листьям), а применение модели для прогнозирования (вычисление значений на заданных кортежах) — «снизу вверх». Существенным отличием двух алгоритмов при этом является то, что в случае R-алгоритма дерево модели T заранее задано и соответствует заранее выбранной кластеризации (или разбиению) множества обучающих данных. В случае L-алгоритма иерархическое разбиение множества данных и

построение дерева T осуществляются в процессе идентификации, при этом точность аппроксимации итоговой иерархической модели в L-алгоритме достигается за счет сложности возникающих в процессе идентификации элементов иерархического разбиения.

В качестве метода анализа остаточных данных при рекуррентной идентификации иерархических моделей с помощью R- и L- алгоритмов предлагается использовать коррелограмму автокорреляционной функции (АКФ) остатков моделей очередного уровня, при этом линейный порядок на множестве этих остатков соответствует упорядочиванию по возрастанию (или убыванию) выходной величины, то есть таргета исходных данных. АКФ остатков иерархических моделей предлагается измерять каждый раз при переходе к следующему уровню иерархии, то есть последовательно для всех промежуточных иерархических моделей. С увеличением уровня иерархии (ростом длины дерева T) величина автокорреляций снижается. Такая ситуация означает, что иерархические модели с добавлением новых уровней точнее аппроксимируют моделируемый процесс, оставляя в остатках все меньше информации. Отсутствие какой-либо информации о процессе в остатках модели очередного уровня эквивалентно ситуации, когда АКФ для остатков мало отличима от АКФ для белого шума. Дальнейшее уточнение иерархической модели в этом случае возможно, но не имеет смысла, поскольку это уточнение равносильно аппроксимации белого шума и приводит к получению переобученной модели. Такая модель не будет обладать необходимым качеством прогнозирования. Таким образом, при использовании разработанных рекуррентных алгоритмов приближение АКФ остатков модели очередного уровня к белому шуму можно считать индикатором останова алгоритмов для исключения дальнейшей переобученности.

В случае построения окрестностных моделей сложных распределенных производственных объектов нередко встречается ситуация, когда имеется несколько технологически разрешенных (номинальных) режимов, зависящих, например, от кластеров входных данных. Соответствующая окрестностная система должна быть мультимодальной, то есть должна описывать все номинальные режимы производственного объекта как «локальные моды» окрестностной системы с возможностью переходов между ними. Соответственно, каждой локальной моде мультимодальной системы отвечает некоторый набор «микрлокальных» мод узлов окрестностной структуры, моделирующей распределенный объект. Предлагаются две схемы построения мультимодальной окрестностной системы: на основе идентификации локальных мод системы в целом (с последующим агрегированием этих мод в общую систему) и на основе агрегирования микрлокальных мод узлов, без идентификации локальных мод всей системы.

В четвертой главе рассматриваются примеры построения иерархических моделей с помощью статического и динамического алгоритмов рекуррентной идентификации с анализом остаточных данных, а также применение алгоритмов к задачам построения иерархических окрестностных моделей прогнозирования модульных характеристик клинкера после обжига и

прогнозирования температурных режимов стадии диффузии производства сахара. Все вычисления проведены на основе использования разработанного комплекса программ, интегрирующего алгоритмы рекуррентной идентификации иерархических окрестностных моделей, метод контроля количества уровней иерархии (длины дерева модели) и модуль вычисления ближайшего обучающего кортежа для входных данных на этапе применения модели для решения задачи прогнозирования. Модули комплекса разработаны на основе программных средств пакетов Matlab, Mathcad и Statistica.

Проведена идентификация традиционных линейной и квадратичной регрессионных моделей, а также иерархических моделей, в последнем случае с помощью L-алгоритма с кластеризацией и трихотомией, для зависимости модульных характеристик клинкера после обжига от химического состава сырьевой муки до обжига. Модульными характеристиками сырьевой муки и клинкера являются представленные в литературе по цементным производствам глиноземный модуль p , силикатный модуль n и коэффициент насыщения KH . Кроме того, для модульных характеристик клинкера построены нейросетевая модель и ансамблевые модели бустинга и бэггинга (случайный лес) на основе деревьев регрессии.

Оценки среднеквадратичной ошибки RMSE и средней относительной ошибки аппроксимации \bar{A} при идентификации различных моделей приведены в таблице 1.

Таблица 1.

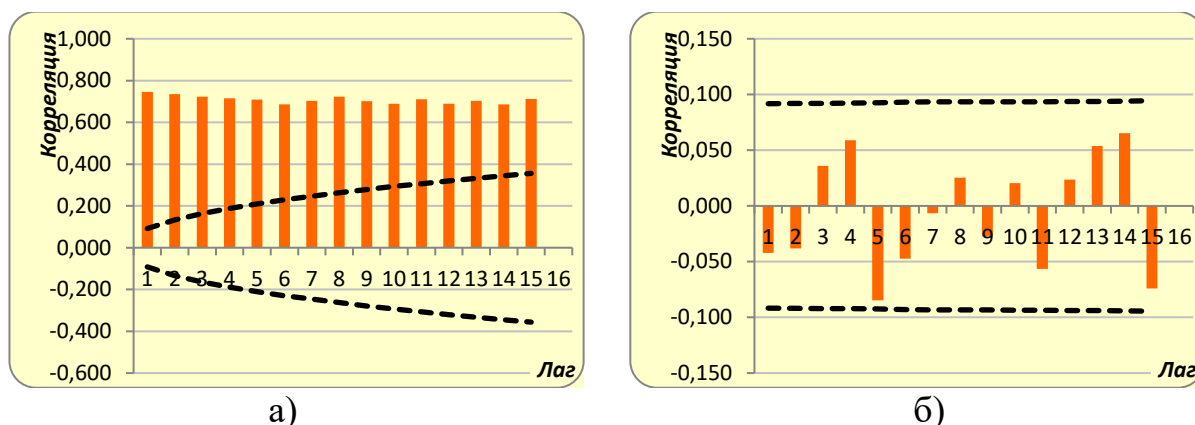
Итоги идентификации различных моделей
модульных характеристик клинкера

Модель	Модульные характеристики					
	p		n		KH	
	RMSE	$\bar{A},\%$	RMSE	$\bar{A},\%$	RMSE	$\bar{A},\%$
Линейная регрессия	0,0478	3,11	0,05	1,64	0,0172	1,39
Квадратичная регрессия	0,047	3,05	0,049	1,61	0,01717	1,39
Бустинг деревьев	0,046	2,92	0,048	1,56	0,01183	0,98
Случайный лес	0,0464	2,98	0,05	1,61	0,0153	1,14
Нейросеть	0,04431	2,84	0,0442	1,43	0,0096	0,81
L-модель 2-уровневая с кластеризацией	0,021	1,33	0,0251	1,02	0,0094	0,75
L-модель 3-уровневая с кластеризацией	0,0075	0,46	0,0085	0,28	0,00436	0,22
L-модель 3-уровневая с трихотомией	0,012	0,68	0,013	0,37	0,0056	0,36
L-модель 4-уровневая с трихотомией	0,0068	0,42	0,0068	0,22	0,002	0,15

Результаты показывают значительное улучшение показателей при моделировании по L-алгоритму рекуррентной идентификации по сравнению с обычными линейной, квадратичной моделями регрессии и по сравнению с ансамблевыми методами деревьев регрессии и нейросетями, обученными в пакете Statistica. Из таблицы 1 видно, что наименьшую ошибку имеет 4-уровневая модель, построенная по L-алгоритму с трихотомией остатков (в таблице выделены полужирным шрифтом).

На рис. 5 показаны коррелограммы АКФ (для лагов 1,...,15) для остатков первоначальной линейной модели регрессии и 4-уровневой модели с

трихотомией, упорядоченных в соответствии с возрастанием выходной величины коэффициента насыщения KH .



а) б)
Рисунок 5 – Изменение коррелограмм остатков KH от линейной регрессии (а) к 4-уровневой модели (б)

Из рис. 5 видно, что автокорреляции остатков существенно снизились после идентификации иерархической модели. При этом близость АКФ остатков иерархической модели к АКФ распределения «белого шума» свидетельствует о достаточной степени извлечения информации моделью из измеренных данных, делается вывод об остановке L-алгоритма на идентификации 4-уровневой модели.

Рассмотрена задача улучшения линейной регрессионной модели зависимости температуры диффузионного сока в нижней части диффузионной колонны от пяти переменных. С этой целью были обучены деревья регрессии и нейросети в пакете Statistica и идентифицирована иерархическая модель по L-алгоритму с трихотомией. Результаты сравнения (оценок среднеквадратичной ошибки RMSE, средней относительной ошибки аппроксимации \bar{A} , коэффициента детерминации R^2) всех моделей представлены в таблице 2.

Таблица 2.

Итоги идентификации различных моделей температуры диффузионного сока в диффузионной колонне

Модель	RMSE	\bar{A} , %	R^2
Линейная регрессия	0,63	0,58	0,64
Бустинг деревьев	0,605	0,58	0,656
Нейросеть	0,556	0,53	0,706
L-модель 3-уровневая с трихотомией	0,277	0,18	0,922

В таблице 2 наилучшие показатели у иерархической модели, построенной по L-алгоритму с трихотомией невязок. Из таблицы видно, что показатель RMSE снизился более чем в 2 раза (примерно на 50%) для 3-уровневой модели по сравнению с линейной моделью, а относительная ошибка \bar{A} снизилась более чем в 3 раза (примерно на 69 %).

Анализ АКФ остатков линейной регрессии, иерархической 2-хуровневой и иерархической 3-хуровневой также показывает последовательное снижение

автокорреляций остатков, упорядоченных в соответствии с возрастанием выходной величины. Остановка L-алгоритма в данной задаче в силу анализа показателей аппроксимации и остаточных данных предложена на идентификации 3-уровневой модели.

Дальнейшая разработка темы диссертации возможна в направлении усовершенствования предложенного метода контроля количества уровней иерархии в моделях и разработки альтернативных методов контроля.

ЗАКЛЮЧЕНИЕ

В работе введены иерархические окрестностные структуры и иерархические окрестностные модели. Разработаны статический и динамический алгоритмы рекуррентной идентификации иерархических окрестностных моделей. Представлена схема анализа остаточных данных промежуточных моделей в алгоритмах рекуррентной идентификации, позволяющая контролировать количество уровней иерархии. Данные алгоритмы и схема анализа остаточных данных были применены для идентификации моделей прогнозирования модульных характеристик клинкера на цементном производстве и прогнозирования температурного режима стадии диффузии производства сахара.

В рамках работы были получены следующие основные результаты:

1. Разработан и исследован новый класс окрестностных моделей, отличающийся иерархической структурой с двусторонними связями и возможностью добавления новых узлов, что позволяет расширить возможности применения окрестностных систем в задачах моделирования производственных процессов и задачах описания иерархических алгоритмов.

2. Разработан алгоритм рекуррентной идентификации иерархических окрестностных моделей на основе аппроксимации невязок промежуточных моделей, отличающийся использованием заданной иерархической кластеризации обучающих входных данных и позволяющий повысить точность моделей (R-алгоритм).

3. Разработан алгоритм и численный метод рекуррентной идентификации иерархических окрестностных моделей на основе трихотомии невязок промежуточных моделей, отличающийся построением в процессе идентификации иерархического разбиения обучающих входных данных и позволяющий улучшать точность моделей и качество прогноза (L-алгоритм).

4. Предложен метод анализа остаточных данных промежуточных моделей в алгоритмах рекуррентной идентификации, отличающийся введением порядка на множестве обучающих данных и позволяющий контролировать количество уровней иерархии, необходимое для достижения соответствия между точностью и качеством прогноза иерархической модели.

5. Разработан комплекс проблемно-ориентированных программ, реализующих алгоритмы рекуррентной идентификации иерархических моделей, отличающихся наличием модуля вычисления ближайшего обучающего кортежа и позволяющих оценивать длину дерева модели.

СПИСОК РАБОТ ПО ТЕМЕ ДИССЕРТАЦИИ

Публикации в изданиях, рекомендованных ВАК РФ

1. Щербаков, А.П. Иерархическая квазилинейная модель прогнозирования свойств клинкера / А.П. Щербаков // Вестник Воронежского государственного технического университета. – 2023. – Т. 19. № 3. – С. 17-22.

2. Мишачев, Н.М. Две схемы иерархической идентификации квазилинейных моделей / Н.М. Мишачев, А.М. Шмырин, А.П. Щербаков // Вестник Воронежского государственного технического университета. – 2023. – Т. 19. № 1. – С. 7-13.

3. Щербаков, А.П. Идентификация мультимодальных окрестностных систем / А.П. Щербаков, А.М. Шмырин, Н.М. Мишачев // Системы управления и информационные технологии. – 2022. – № 3 (89). – С. 24-28.

4. Шмырин, А.М. Два подхода к исследованию общего параметрического уравнения окрестностной модели печи обжига клинкера / А.М. Шмырин, И.А. Седых, А.П. Щербаков, А.Г. Ярцев // Системы управления и информационные технологии. – 2015. – №1-1 (59). – С. 185-189.

5. Шмырин, А.М. Общие билинейные дискретные модели / А.М. Шмырин, И.А. Седых, А.П. Щербаков // Вестник Воронежского государственного технического университета. – 2014. – Т. 10. № 3-1. – С. 44-49.

Публикации, индексируемые в Scopus и Web of Science

6. Mishachev, N.M. Approximation of curves in phase space by solutions of control system / N.M. Mishachev, A.M. Shmyrin, A.P. Shcherbakov // 2022 4rd International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency, SUMMA 2022. – 2022. – P. 46-48.

7. Mishachev, N.M. Polymodal neighborhood systems in system modeling course / N.M. Mishachev, A.M. Shmyrin, A.P. Shcherbakov // 2022 2nd International Conference on Technology Enhanced Learning in Higher Education, TELE 2022. – 2022. – P. 152-155.

Свидетельства на программы для электронных вычислительных машин

8. Свидетельство о государственной регистрации программы для ЭВМ RU 2015614243 Российская Федерация. Расчет корреляционных размерностей и доли случайного хаоса по данным временного ряда : № 2015610763 : зарег. 13.02.2015 : опубл. 10.04.2015 / А.М. Шмырин, А.П. Щербаков, И.А. Седых, А.Г. Ярцев : правообл. Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Липецкий государственный технический университет» (ФГБОУ ВПО ЛГТУ). – Версия операционной системы: Windows XP/Vista/7/8/10. – Объем: 0,6 Мб.

9. Свидетельство о государственной регистрации программы для ЭВМ RU 2015610766. Российская Федерация. Получение общего параметрического уравнения окрестностной модели : № 2014661696 : зарег. 19.11.2014 : опубл. 16.01.2015 / А.Г. Ярцев, А.М. Шмырин, И.А. Седых, А.П. Щербаков : правообл. Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Липецкий

государственный технический университет» (ФГБОУ ВПО ЛГТУ) – Версия операционной системы: Windows XP. – Объем: 1,1 Мб.

10. Свидетельство о государственной регистрации программы для ЭВМ RU 2015617475 Российская Федерация Исследование и решение систем методом ортонормализации Грама-Шмидта : № 2015612504 : заявл. 01.04.2015 : опубл. 10.07.2015 / В.В. Семёнова, А.М. Шмырин, И.А. Седых, А.П. Щербаков, А.Г. Ярцев : правообл. Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Липецкий государственный технический университет» (ФГБОУ ВПО ЛГТУ) – Версия операционной системы: Windows XP и выше. – Объем: 88,5 Кб.

Статьи и материалы конференций

11. Щербаков, А.П. L-алгоритм моделирования переменных состояния стадии диффузии производства сахара / А.П. Щербаков // Вестник Липецкого государственного технического университета. – 2023. – № 3 (52). – С. 21-27.

12. Щербаков, А.П. Трехуровневая иерархическая регрессионная модель / А.П. Щербаков, А.М. Шмырин, Н.М. Мишачев // Нано-био-технологии. Теплоэнергетика. Математическое моделирование. Сборник статей международной научно-практической конференции. – Липецк, 2023. – С. 283-288.

13. Щербаков, А.П. Генерирование тестовых данных для регрессионной идентификации квазилинейных иерархических моделей / А.П. Щербаков // Вестник Липецкого государственного технического университета. – 2022. – № 3 (49). – С. 41-47.

14. Щербаков, А.П. Представление нейронных сетей в виде общих окрестностных моделей / А.П. Щербаков // Актуальные проблемы естественных наук и их преподавания. Материалы научной конференции молодых ученых. – Липецк, 2016. – С. 282-288.

15. Шмырин, А.М. Наличие экстремумов параметрического уравнения печи обжига клинкера / А.М. Шмырин, И.А. Седых, А.П. Щербаков, А.Г. Ярцев // Вести высших учебных заведений Черноземья. – 2015. – №1 (39). – С. 62-67.

16. Шмырин, А.М. Исследование окрестностной модели печи обжига клинкера с учетом допустимых значений параметров / А.М. Шмырин, И.А. Седых, А.П. Щербаков, А.Г. Ярцев // Вестник Липецкого государственного технического университета. – 2015. – №2 (24). – С. 11-14.

17. Shmyrin, A.M. Research trilinear neighborhood model of clinker kiln / A.M. Shmyrin, I.A. Sedykh, A.P. Shcherbakov, A.G. Yartsev // Modern informatization problems in simulation and social technologies: Proceedings of the XX-th International Open Science Conference (Yelm, WA, USA). – 2015. – P. 202-206.

18. Щербаков А.П. К вопросу исследования хаотичности входного процесса / А.П. Щербаков // Школа молодых учёных по проблемам гуманитарных, естественных, технических наук. Сборник материалов областного профильного семинара. – Елец, 2014. – С. 234-237.

19. Шмырин, А.М. Методы нелинейного анализа при исследовании характеристик производства клинкера / А.М. Шмырин, И.А. Седых,

А.П. Щербаков // Вестник Тамбовского университета. Серия: Естественные и технические науки. – 2014. – Т. 19 № 3. – С. 923-926.

20. Шмырин, А.М. Параметрическое окрестностное моделирование печи обжига клинкера / А.М. Шмырин, И.А. Седых, А.П. Щербаков, А.Г. Ярцев, Е.С. Аникеев // Вестник Тамбовского университета. Серия: Естественные и технические науки. – 2014. – Т. 19 № 3. – С. 927-930.

21. Шмырин, А.М. Параметрическая билинейная окрестностная модель системы теплоснабжения / А.М. Шмырин, А.П. Щербаков, А.Г. Ярцев // Современные методы прикладной математики, теории управления и компьютерных технологий: сб. тр. VI междунар. конф. «ПМТУКТ-2013». – 2013. – С. 277-279.

22. Шмырин, А.М. Окрестностное моделирование двумерных нелинейных динамических систем / А.М. Шмырин, И.А. Седых, В.М. Тюрин, В.Б. Васильев, А.П. Щербаков // Вестник Тамбовского университета. Серия: Естественные и технические науки. – 2013. – Т. 18 № 1. – С. 81-88.

23. Шмырин, А.М. Исследование поведения окрестностной системы с учетом допустимых значений параметров / А.М. Шмырин, А.П. Щербаков, А.Г. Ярцев // Вестник Тамбовского университета. Серия: Естественные и технические науки. – 2013. – Т. 18 № 5-2. – С. 2747-2748.

24. Шмырин, А.М. Окрестностное моделирование распределенных нелинейных динамических систем / А.М. Шмырин, С.С. Роечко, А.П. Щербаков // Управление большими системами. Материалы IX Всероссийской школы-конференции молодых ученых. – Тамбов, 2012. – Т. 1. – С. 104-107.

25. Шмырин, А.М. Окрестностный подход к моделированию распределенных динамических систем / А.М. Шмырин, В.В. Кавыгин, С.С. Роечко, А.П. Щербаков // Фундаментальные и прикладные проблемы модернизации современного машиностроения и металлургии. Сборник научных трудов международной научно-технической конференции, посвященной 50-летию кафедры технологии машиностроения ЛГТУ. – Липецк, 2012. – Том Часть 2 – С. 321-326.

26. Щербаков, А.П. Моделирование логистического отображения билинейными окрестностными системами / А.П. Щербаков, А.М. Шмырин // Современные проблемы информатизации в экономике и обеспечении безопасности: сб. тр. – Воронеж: Научная книга, 2012. – № 17. – С. 119-123.

Подписано в печать 24.01.2024. Формат 60×84 1/16. Бумага офсетная.

Цифровая печать. Объем 1,2 п.л. Тираж 100 экз. Заказ № 13.

Липецкий государственный технический университет.

Отдел полиграфических работ и услуг.

398055, Липецк, ул. Московская, 30.