

Evaluation of Operating Modes for Central Heating Stations Using Neural Network and Ensemble Machine Learning Methods

Dvortsevov A.I., Borush O.V., Yakovina I.N., Boyko E.E., Vindemut D.Y., Yankov P.A.

Novosibirsk State Technical University
Novosibirsk, Russian Federation

Abstract. The main objective of this study is to develop a method for assessing the operating modes of central heating stations (CHS) based on forecasting the thermal power and analyzing the deviations of actual parameters from the reference values. To achieve this goal, the following tasks were solved: primary processing and cleaning of data on technological parameters of CHS and meteorological observations for the period 2022-2024 were performed; regression models (Random Forest, XGBoost, CatBoost and a fully connected neural network) were trained and compared for forecasting the thermal power; an algorithm for assessing and forecasting the operating modes of CHS was developed. The most important results are: the fully connected neural network model, which achieved the determination coefficient $R^2 = 0.9997$ and the mean absolute error $MAE = 0.0099$ Gcal/h; An algorithm developed based on the resulting model allows for differentiating deviations in central heating system operating modes by their source—from the heat supplier, from the consumer, and hidden abnormal conditions not detected by direct comparison with the temperature chart. The analysis revealed that a sharp increase in outside air temperature leads to an increase in the direct network water pressure due to the specifics of quantitative weather-dependent regulation, which can cause accelerated wear of heating networks. The significance of the obtained results lies in the fact that the proposed algorithm ensures the early detection of emergency and pre-emergency conditions, which helps reduce accident rates, reduce repair and maintenance costs, and improve the energy efficiency of district heating systems. The method can be adapted to various regions with centralized heating.

Keywords: district heating substation, machine learning, neural network forecasting, regression analysis, anomaly detection, thermal power, temperature schedule, weather-dependent control, pre-emergency mode, heat supply monitoring.

UDC: 697.34

DOI: <https://doi.org/10.52254/1857-0070.2026.3-71.14>

Evaluarea modurilor de funcționare ale stațiilor centrale de încălzire prin utilizarea metodelor de învățare automată bazate pe rețele neuronale și pe ansambluri

Dvorțevov A.I., Boruș O.V., Yakovina I.N., Boiko E.E., Vindemut D.Y., Iankov P.A.

Universitatea Tehnică de Stat din Novosibirsk, Novosibirsk, Federația Rusă

Rezumat. Obiectivul principal al acestui studiu este dezvoltarea unei metode de evaluare a regimurilor de funcționare ale centralelor termice (CTC), bazată pe prognozarea puterii termice și pe analiza abaterilor parametrilor reali față de valorile de referință. Pentru atingerea acestui obiectiv au fost realizate următoarele sarcini: prelucrarea primară și curățarea datelor privind parametrii tehnologici ai CTC și a observațiilor meteorologice pentru perioada 2022–2024; antrenarea și compararea unor modele de regresie (Random Forest, XGBoost, CatBoost și o rețea neuronală complet conectată) pentru prognozarea puterii termice; dezvoltarea unui algoritmul pentru evaluarea și prognozarea regimurilor de funcționare ale CTC. Cele mai importante rezultate sunt următoarele: modelul de rețea neuronală complet conectată a obținut un coeficient de determinare $R^2 = 0,9997$ și o eroare medie absolută $MAE = 0,0099$ Gcal/h. Algoritmul dezvoltat pe baza acestui model permite diferențierea abaterilor regimurilor de funcționare ale sistemului de încălzire centralizată în funcție de sursa acestora – furnizorul de energie termică, consumatorul sau condițiile anormale ascunse, care nu pot fi detectate prin compararea directă cu diagrama de temperatură. Analiza a evidențiat faptul că o creștere bruscă a temperaturii aerului exterior conduce la majorarea presiunii apei pe conducta de tur a rețelei termice, ca urmare a particularităților reglării cantitative dependente de condițiile meteorologice, ceea ce poate accelera uzura rețelelor de încălzire. Semnificația rezultatelor obținute constă în faptul că algoritmul propus asigură detectarea timpurie a situațiilor de urgență și premurgătoare acestora, contribuind astfel la reducerea frecvenței avariilor, la diminuarea costurilor de reparație și întreținere și la creșterea eficienței energetice a sistemelor de încălzire centralizată. Metoda propusă poate fi adaptată pentru diverse regiuni în care sunt utilizate sisteme de încălzire centralizată.

Cuvinte-cheie: substație de termoficare, învățare automată, prognoză prin rețele neuronale, analiză de regresie, detectarea anomaliilor, energie termică, program de temperatură.

Оценка режимов работы центральных тепловых пунктов с применением нейросетевых и ансамблевых методов машинного обучения

Дворцовой А.И., Боруш О.В., Яковина И.Н., Бойко Е.Е., Виндемут Д.Ю., Яньков П.А.

Новосибирский государственный технический университет, Новосибирск, Российская Федерация

Аннотация. Основной целью настоящего исследования является разработка метода оценки режимов работы центральных тепловых пунктов (ЦТП) на основе прогнозирования тепловой мощности и анализа отклонений фактических параметров от эталонных значений. Для достижения поставленной цели были решены следующие задачи: выполнена первичная обработка и очистка данных технологических параметров ЦТП и метеорологических наблюдений за период 2022–2024 гг.; обучены и сравнены модели регрессии (Random Forest, XGBoost, CatBoost и полносвязная нейросеть) для прогнозирования тепловой мощности; разработан алгоритм оценки режимов работы ЦТП. Наиболее важными результатами являются: полносвязная нейросетевая модель, которая достигла коэффициента детерминации $R^2 = 0.9997$ и средней абсолютной ошибки $MAE = 0.0099$ Гкал/ч; разработанный на основе полученной модели алгоритм, позволяющий дифференцировать отклонения в режимах работы ЦТП по источникам возникновения — со стороны поставщика тепла, со стороны потребителя и скрытые нештатные режимы, не фиксируемые прямым сравнением с температурным графиком. В ходе анализа было установлено, что резкое повышение температуры наружного воздуха приводит к росту давления прямой сетевой воды из-за особенностей количественного погодозависимого регулирования, что может вызывать ускоренный износ тепловых сетей. Значимость полученных результатов состоит в том, что предложенный алгоритм обеспечивает раннее обнаружение аварийных и предаварийных режимов, что способствует снижению аварийности, уменьшению затрат на ремонт и обслуживание, а также повышению энергоэффективности систем централизованного теплоснабжения. Метод может быть адаптирован для различных регионов с централизованным теплоснабжением.

Ключевые слова: центральный тепловой пункт, машинное обучение, нейросетевое прогнозирование, регрессионный анализ, обнаружение аномалий, тепловая мощность, температурный график, погодозависимое регулирование, предаварийный режим, мониторинг теплоснабжения.

ВВЕДЕНИЕ

Исследования в области систем централизованного теплоснабжения (СЦТ) можно условно разделить на три направления.

Первое направление — работы, посвящённые эксплуатационным и нормативно-регуляторным аспектам систем теплоснабжения.

Второе направление — исследования по прогнозированию тепловой нагрузки.

Третье направление — работы по обнаружению нештатных режимов работы.

Важнейшим звеном СЦТ являются центральные тепловые пункты (ЦТП), обеспечивающие подачу теплоносителя потребителям. Традиционное регулирование температуры прямой сетевой воды по фиксированному температурному графику (например, 95/70 °C) имеет существенный недостаток: отсутствие возможности оперативно учитывать меняющийся тепловой баланс зданий, обусловленный солнечной радиацией, ветром и внутренними тепловыделениями [1, 2]. Как следствие, возникают систематические «перетопы», ведущие к перерасходу топливно-энергетических ресурсов, и «недотопы», нарушающие тепловой комфорт и

повышающие риск аварий [3]. Это диктует необходимость перехода к интеллектуальным системам мониторинга, которые позволяют прогнозировать потребление и выявлять аномалии с возможностью разделения ответственности за отклонения.

В мировой практике машинное обучение активно проникает в теплоэнергетику. Обзорные работы [4–6] свидетельствуют о формировании двух основных направлений: прогнозирования тепловой нагрузки и обнаружения неисправностей. Методы машинного обучения, применяемые в теплоснабжении, тепловых сетях и на тепловых пунктах, охватывают широкий спектр алгоритмов — от классических ансамблевых моделей до глубоких нейросетевых архитектур. Для прогнозирования теплопотребления успешно используются случайный лес (RandomForest), градиентный бустинг (XGBoost, CatBoost), а также полносвязные и рекуррентные нейронные сети (LSTM), а также ансамблевые алгоритмы [7–14, 22–24]. Несмотря на достигнутый прогресс, анализ публикаций [4, 5, 22, 24, 27] выявляет несколько нерешённых проблем, ограничивающих внедрение машинного обучения на данных реальных

ЦТП. Во-первых, стоит выделить отсутствие комплексных решений – большинство работ реализует либо прогноз, либо диагностику, но не то и другое одновременно. Во-вторых, потребность в размеченных данных – эффективные диагностические модели часто требуют исторических примеров аварий [5, 12], которые редко доступны на практике. В-третьих, недостаточный учёт цикличности теплопотребления [29, 30].

Наконец, существующие подходы не способны дифференцировать источники отклонений, поскольку не разделяют сбои, вызванные поставщиком тепла (нарушение температурного графика на вводе), изменениями на стороне потребителя (изменение водоразбора, самовольная реконструкция) и скрытыми нештатными режимами.

Научная новизна исследования заключается в создании комплексного алгоритма оценки режимов работы ЦТП, который, в отличие от известных подходов, одновременно обеспечивает: прогнозирование тепловой мощности и выявление аномалий без привлечения размеченных аварийных выборок, использует полносвязную нейросеть с циклическими календарными признаками и позволяет дифференцировать источники отклонений (поставщик тепла, потребитель, скрытые нештатные режимы) на основе двухуровневого контроля температурного графика и результатов анализа оценок, полученных с помощью прогнозной модели.

Цель настоящей работы — разработка и экспериментальная верификация алгоритма оценки и прогнозирования режимов работы ЦТП, обеспечивающего раннее обнаружение предаварийных ситуаций

Для достижения поставленной цели решались следующие задачи:

1. Выполнить первичную обработку и формирование набора признаков на основе данных технологических и метеорологических измерений за 2022-2024 годы.
2. Провести сравнительный анализ ансамблевых алгоритмов (Random Forest, XGBoost, CatBoost) и полносвязной нейронной сети для прогнозирования потребления тепловой энергии Q для последующей оценки режимов работы.
3. Разработать алгоритм оценки режимов работы ЦТП на основе анализа температур

подающего и обратного трубопроводов с утверждённым температурным графиком (с доверительным интервалом $\pm 3^\circ\text{C}$) и выявления аномалий в работе ЦТП.

4. Экспериментально оценить эффективность разработанного алгоритма на реальных данных и сформулировать рекомендации по его внедрению в системы мониторинга.

2. МЕТОД

Для реализации поставленных целей был проведен сбор и подготовка обширного комплекса данных, включающего технологические параметры работы ЦТП, метеорологические наблюдения и нормативные требования. Эти данные позволили сформировать основу для последующего моделирования и анализа режимов работы тепловых станций.

Первую группу составляют ключевые технологические параметры функционирования тепломеханического оборудования ЦТП. К ним относятся термодинамические показатели теплоносителя на входе и выходе из системы: температура ($T_1, T_2, ^\circ\text{C}$) и давление ($P_1, P_2, \text{кгс/см}^2$). Для учета материального баланса фиксируются массовые расходы теплоносителя в подающем ($M_1, M_2, \text{т/ч}$) трубопроводах.

Потребление тепловой энергии, рассчитываемое на основе разности энтальпий потоков ($Q, \text{Гкал}$). Принципиальная схема ЦТП с двухступенчатым смешанным подключением подогревателей ГВС и независимым присоединением отопительной нагрузки представлена на рисунке 1.

Вторая группа включает среднечасовые метеорологические параметры, определяющие динамику теплопотерь зданий и величину тепловой нагрузки.

Вектор внешних факторов формируется значениями температуры наружного воздуха ($T_{nv}, ^\circ\text{C}$), скорости ветра ($V_{avg}, \text{м/с}$), относительной влажности ($Wl, \%$) и атмосферного давления ($P_{atm}, \text{гПа}$).

Третьим компонентом набора данных является нормативно-технический базис — утвержденный эталонный температурный график ЦТП (в исследуемом контуре рассмотрены режимы 150/70, 130/70 и 115/70). Данный график задает функциональную зависимость целевых температур в подающем и обратном трубопроводах от среднесуточной температуры наружного воздуха, выступая в

качестве референсных значений при анализе аномалий и отклонений.

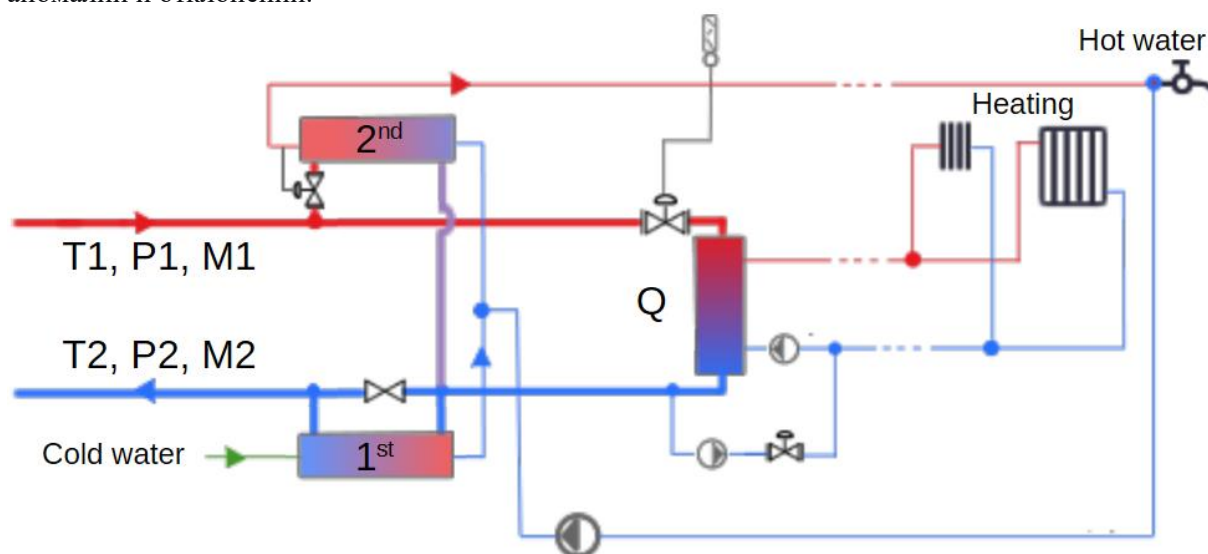


Рис. 1. Принципиальная схема ЦТП с двухступенчатым смешанным подключением подогревателей ГВС и независимым присоединением отопительной нагрузки.
Fig1. Principal diagram of the CHS system with a two-stage mixed connection of heaters and an independent connection of heating.

2.1. Первичная обработка данных и анализ выбросов

Исходные данные включают технологические параметры, метеорологические наблюдения и эталонный температурный график. Частота дискретизации — 1 час.

В ходе первичной обработки были удалены явные технические выбросы (например, нулевые или отрицательные значения расходов при работающем оборудовании, скачки давления, превышающие паспортные значения датчиков) и заполнены пропуски в исходных данных методом линейной интерполяции для пропусков длительностью до 3 часов и переносом предыдущего значения (forward fill) для пропусков более 3 часов (что обосновано инерционностью тепловых процессов).

Аномальные периоды не удалялись автоматически, а анализировались с точки зрения физической допустимости и сопоставлялись с предоставленным списком аварий и плановых ремонтных работ ЦТП. Это позволило сохранить в выборке корректные, но нестандартные режимы, которые важны для обучения моделей обнаружения аномалий.

Фрагмент исходных данных отражающих сложный характер моделируемых процессов приведен на рис. 2. Представленный фрагмент выборки иллюстрирует выраженную сезонную динамику тепловых режимов:

максимальные значения $T1$ и $T2$ фиксируются в зимние месяцы (январь, март, ноябрь), а минимальные — в летний период (июль–август). При этом разность $\Delta T = T1 - T2$, косвенно характеризующая теплопередачу к потребителю, варьируется от 2–3 °C летом до 20–30 °C зимой, что соответствует изменению тепловой нагрузки в течение года. Такая межгодовая изменчивость, наряду с нелинейными перепадами между соседними месяцами (например, резкий рост $T1$ с 72 °C в июле до 75 °C в августе 2023 г., а затем скачок до 88 °C в январе 2024 г.), подтверждает сложный, многомерный характер моделируемых теплогидравлических процессов. Данные не подчиняются простым линейным закономерностям и содержат скрытые зависимости, обусловленные погодными факторами, режимом работы оборудования и сезонными пиками потребления. Это обосновывает целесообразность применения методов машинного обучения для построения прогнозных моделей тепловой мощности и оптимизации управления центральным тепловым пунктом.

Для оценки взаимосвязей между признаками на первом этапе анализа были построены корреляционные матрицы методами Пирсона и Кендалла. Установлено, что потребление тепловой энергии Q наиболее сильно коррелирует с $T1$ (0.71), $M1$ (0.85), $P2$

(0.61) и M2 (0.82). Однако, отказ от включения в состав разрабатываемых моделей этих параметров приводил к существенному

снижению точности (более 15%), поэтому они были включены в результирующее признаковое пространство.

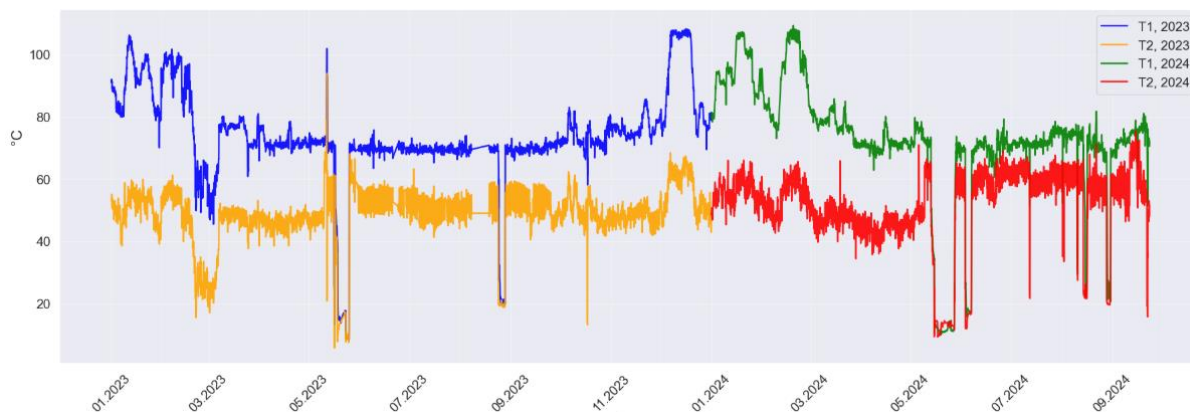


Рис 2. Фрагмент исходных данных объекта теплоснабжения (температура подающего и обратного теплоносителя T1, T2).

Fig. 2. Fragment of source data of the central heating point (the temperature of the supply and return water T1, T2).

2.2. Формирование расширенного набора признаков

Для повышения точности прогнозирования исходные данные были дополнены тремя категориями производных признаков.

Категория 1. Календарные и циклические признаки. На основе временной метки был сформирован набор признаков: час суток, день недели, месяц года, индикатор выходного/праздничного дня, индикатор отопительного периода. Применены синусно-косинусные преобразования для кодирования циклических признаков. Это позволяет улучшить качество прогнозирующих моделей в задачах, связанных с временными рядами.

$$x_{sin} = \sin\left(2\pi \frac{x}{T}\right), x_{cos} = \cos\left(2\pi \frac{x}{T}\right) \quad (1)$$

где T — длина периода (24 для часа, 12 для месяца, 7 для дня недели). Дополнительно введены бинарные сезонные флаги: зима, весна, лето, осень и непрерывные сезонные признаки на основе дня года.

Традиционные методы кодирования временных признаков, такие как простое присвоение чисел часам суток или дням недели, приводят к некорректному представлению цикличности, поскольку искусственно создают разрыв между соседними значениями (например, между 23:00 и 00:00 или декабрем и январем).

Категория 2. Лаговые признаки и скользящие средние. Для учёта тепловой инерции зданий и сетей выполнен отбор информативных лагов температуры

наружного воздуха методом взаимной информации. Выбраны пять наиболее значимых лагов: 8, 10, 12, 16 и 20 часов, что соответствует физическому времени запаздывания реакции системы отопления на изменение внешней температуры (от 8 часов для малых зданий до 20 часов для крупных микрорайонов с развитой теплосетью). Для каждого из этих лагов рассчитаны скользящие средние и трендовые признаки (разность между текущим значением и значением 3 и 6 часов назад), что позволило моделям учитывать наряду с исходными значениями динамику изменения погодных условий.

Категория 3. Производные теплотехнические признаки. К исходным признакам характеризующим гидравлический режим (P1, P2, M1) была добавлена разность температур подающего и обратного трубопроводов $dT = T1 - T2$.

2.3. Итоговый набор признаков и разделение выборки

Финальный набор признаков для прогнозирования тепловой мощности Q включает:

1. Метеорологические (4 признака).
2. Лаги T_{nv} (8, 10, 12, 16, 20 ч).
3. Скользящие средние T_{nv} (8, 10, 12, 16, 20 ч).
4. Тренды T_{nv} (3 и 6 ч).
5. Циклические временные признаки (8 признаков).
6. Сезонные признаки (4 признака).
7. Теплотехнические признаки: M1, T1, dT.

Все числовые признаки нормированы для алгоритмов, чувствительных к масштабу.

3. РАЗРАБОТКА МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ

3.1. Общая постановка эксперимента

Представленное в данной работе решение задачи оценки тепловой мощности Q получено с использованием четырех алгоритмов: Random Forest, XGBoost, CatBoost и полносвязной нейронной сети. Все эти модели обучались на едином наборе признаков, описанном в разделе 2. Оценка качества выполнялась по метрикам: средняя абсолютная ошибка (MAE), среднеквадратичная ошибка (RMSE), коэффициент детерминации (R^2) и средняя абсолютная процентная ошибка (MAPE). Обучение проводилось на сегментах исходной выборки из каждого сезона (80 % всего массива данных), тестирование — на оставшихся 20% данных временного ряда с сохранением временной упорядоченности и стратификацией по сезонам с использованием опции *StratifiedShuffleSplit*. Модели тестировались на данных, не участвовавших в обучении, что имитирует реальные условия эксплуатации. Результаты представлены в таблице 1, пояснения по каждой модели приведены ниже.

Анализ распределения ошибок для перечисленных моделей показал, что ошибки распределены равномерно относительно нуля, медианное значение остатков близко к нулю, что указывает на отсутствие систематического смещения. Небольшое увеличение дисперсии ошибок наблюдается на малых значениях Q (менее 0.5 Гкал/ч), что соответствует летнему режиму и объясняется меньшим объемом данных в этом временном диапазоне.

Таблица 1. Сравнение метрик прогнозирования тепловой мощности Q .
Table 1. Comparison of thermal power prediction metrics Q

Модель	MAE, Гкал/ч	RMSE, Гкал/ч	R^2	MAPE, %
Random Forest	0.0225	0.0401	0.9984	13.33
XGBoost	0.0165	0.0279	0.9992	12.05
CatBoost	0.0135	0.0217	0.9995	16.74
Нейросеть (MLP)	0.0099	0.0185	0.9997	10.99

3.2. Результаты обучения моделей

3.2.1. Случайный лес (Random Forest)

Модель реализована с использованием *RandomForestRegressor* из библиотеки *scikit-learn*.

Выбор данного алгоритма обусловлен устойчивостью к переобучению и способностью работать с нелинейными зависимостями без необходимости масштабирования признаков.

Проведенный анализ позволил выявить ключевые факторы, оказывающие наибольшее влияние на точность прогнозирования тепловой мощности. Среди них особенно выделяются скользящее среднее температуры наружного воздуха за последние 20 часов (58.1%), массовый расход сетевой воды (17.9%) и температура подающего трубопровода (12.3%). Полученные результаты подтверждают важность учета климатических условий и гидродинамических характеристик системы при построении прогностических моделей.

3.2.2. Экстремальный градиентный бустинг (XGBoost)

Модель реализована с использованием библиотеки *Xgboost*.

Для предотвращения переобучения использовался механизм ранней остановки.

Анализ важности признаков выявил, что бинарный признак зимы имеет наибольшую значимость (32.9%), что обусловлено резким изменением режима теплоснабжения между отопительным и межотопительным периодами. Скользящее среднее $T_{nv_ma_20h}$ — 27.4%, бинарный признак лета — 12.6%, $M1$ — 5.9%, dT — 4.5%. XGBoost показал лучшее качество, чем Random Forest, по всем метрикам.

3.2.3. Градиентный бустинг CatBoost

Модель реализована с использованием библиотеки *Catboost*.

Полученная модель даёт более точные предсказания в области высоких нагрузок (зима), но допускает большие относительные отклонения в межотопительный период.

Анализ важности признаков показал, что наибольший вклад вносят: dT (разность температур) — 28.8%, $M1$ — 26.5%, $T1$ — 10.8%, а также циклические признаки дня года. Это подтверждает, что для CatBoost сезонная цикличность оказалась более информативной, чем бинарные сезонные флаги, использованные в XGBoost.

3.2.4. Полносвязная нейронная сеть (многослойный перцептрон)

Архитектура нейросетевой модели на базе TensorFlow/Keras включала входной слой (31 признак), четыре полносвязных слоя (Dense) с 256, 128, 64 и 32 нейронами и функцией активации ReLU, а также выходной слой с одним нейроном и линейной активацией.

Нейросетевая модель превзошла все ансамблевые алгоритмы по всем метрикам, что подтверждает её способность аппроксимировать сложные нелинейные зависимости без явного задания структуры модели.

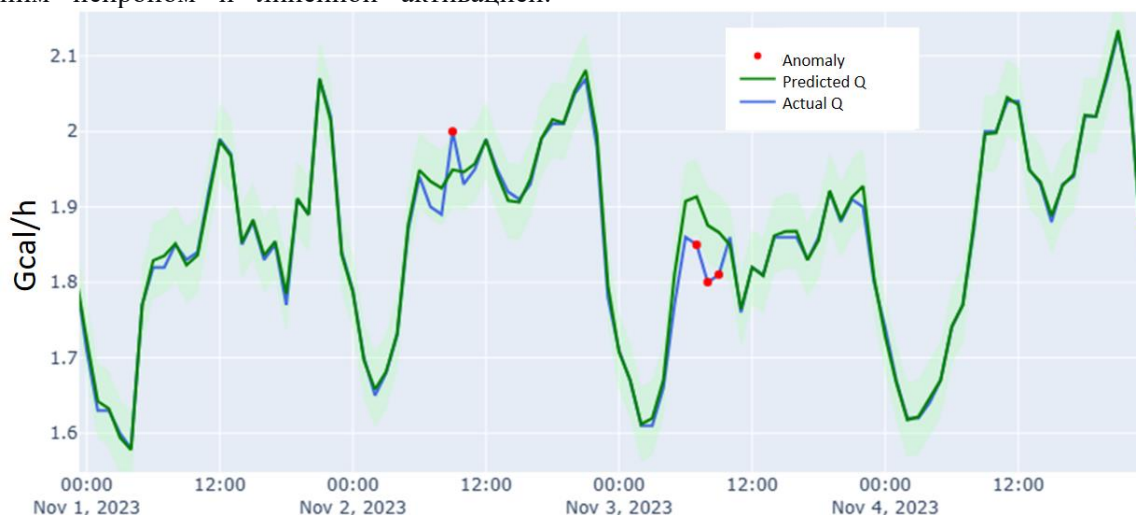


Рис. 3 – Реальные и прогнозируемые значения тепловой мощности нейросетевой моделью.
Fig. 3 – Real and predicted values of thermal power by the neural network model.

На основе анализа полученных результатов (табл. 1, рис. 3) для дальнейшего использования выбрана полносвязная нейронная сеть как модель с наилучшими метриками точности и устойчивостью к нелинейностям.

4. АНАЛИЗ РЕЖИМОВ РАБОТЫ ЦТП

На основе обученной нейросетевой модели (раздел 3.2.4) разработан комплексный алгоритм анализа режимов работы центрального теплового пункта.

Алгоритм включает два независимых шага проверки: сравнение температур подающего и обратного трубопроводов с утверждённым температурным графиком и анализ отклонений тепловой мощности от прогноза нейросетевой модели.

4.1. Эталонный температурный график

Для анализа режимов работы используется утверждённый температурный график (табл. 2), содержащий зависимость температур в подающем и обратном трубопроводе (режимы 150/70, 130/70, 115/70) от среднесуточной температуры наружного воздуха.

Таблица 2. Фрагмент эталонного температурного графика.
Table 2. Fragment of the reference temperature graph.

T _{nv_avg} , °C	R ₁₅₀ , °C	R ₁₃₀ , °C	R ₁₁₅ , °C	T ₂ , °C
+1	70.0	64.5	58.8	42.4
0	72.4	66.5	60.5	43.4
-1	74.7	68.5	62.2	44.2
-2	77.0	70.4	63.8	45.0
-3	79.3	72.4	65.5	45.9

График служит нормативной базой для определения режима работы ЦТП при текущих погодных условиях.

4.2. Анализ режимных параметров T1 и T2

Алгоритм анализа температуры прямой и обратной сетевой воды включает две последовательные проверки. Первая проверка сопоставляет фактическую температуру подающего трубопровода T1 с эталонными значениями для режимов R₁₅₀, R₁₃₀, R₁₁₅ с учётом доверительного интервала $\pm 3^\circ\text{C}$. При T_{nv_avg} 8°C устанавливается режим R₀ (отопление отключено). При совпадении с одним из режимов фиксируется статус «ок»; в

противном случае — статус «an_T1» (отклонение со стороны поставщика тепла). Вторая проверка выполняется только при статусе «ok» для T1: фактическая температура обратного трубопровода T2 сравнивается с эталонным значением из графика. При попадании в интервал $\pm 3^\circ\text{C}$ возвращается статус «ok», иначе — «an_T2» (проблемы у потребителя). Доверительный интервал $\pm 3^\circ\text{C}$ обоснован «Правилами технической эксплуатации тепловых энергоустановок» и подтверждён анализом распределения отклонений на валидационной выборке.

4.3. Анализ тепловой мощности Q

Для Q, в отличие от T1 и T2, не существует утверждённого графика, поэтому в качестве эталона используется прогноз обученной нейросетевой модели. Алгоритм включает: подачу входных признаков на модель, обратное масштабирование предсказания, вычисление абсолютной ошибки (формула 3) и сравнение с порогом 0.05 Гкал/ч. Порог обнаружения аномалий определяется для каждой модели и объекта.

$$\text{error} = |Q_{\text{actual}} - Q_{\text{predicted}}| \quad (3)$$

где Q_{actual} — показания датчика, $Q_{\text{predicted}}$ — предсказание модели.

4.4. Результаты комплексного анализа

Применение алгоритма ко всей тестовой выборке (15 915 записей) дало следующие результаты (табл. 3).

Таблица 3. Статистика отклонений тепловой мощности Q.

Table 3. Statistics of deviationsof thermal power diagnostics Q.

Период	Всего записей	Норма (ok)	Аномалия (an_Q)
Весь период	15 915	15 612 (98.1%)	303 (1.9%)
Отопительный период	8 898	8 852 (99.5%)	46 (0.5%)

Доля аномалий в отопительный период значительно ниже (0.5% против 1.9% в целом), что объясняется стабильной работой системы теплоснабжения при высоких тепловых нагрузках. В межотопительный период точность прогноза снижается из-за малых абсолютных значений Q и более изменчивых режимов.

Фрагмент результатов анализа тепловой мощности Q представлен в табл. 4.

Таблица 4. Фрагмент результатов анализа тепловой мощности Q.

Table 4. Fragment of the results of the analysis of thermal power Q.

Дата и время	Q_факт, Гкал/ч	Q_прогноз, Гкал/ч	O_ошибка, Гкал/ч	Статус
2022-12-14 06:00	3.110	3.118	0.008	ok
2022-12-14 07:00	3.080	3.094	0.014	ok
2022-12-14 08:00	1.680	1.779	0.099	an_Q
2022-12-14 09:00	0.000	-0.077	0.077	an_Q
2022-12-14 10:00	2.070	2.076	0.006	ok
2022-12-14 11:00	3.690	3.352	0.338	an_Q
2022-12-14 12:00	3.390	3.282	0.108	an_Q
2022-12-14 13:00	3.310	3.273	0.037	ok

Как видно из таблицы, в ряде случаев (08:00, 09:00, 11:00, 12:00) ошибка значительно превышает порог 0.05 Гкал/ч. Сопоставление с журналом эксплуатационных событий показало, что в этот период проводились ремонтные работы, что привело к временному снижению теплоснабжения.

Таблица 5. Общая статистика режимов работы ЦТП.

Table 5. General statistics for operation modes of central heating station.

Статус	Количество записей	Доля, %	Интерпретация
ok	9 040	56.8	Все параметры в норме
an_T1	4 794	30.1	Отклонение T1 от графика (поставщик)
an_T2	2 081	13.1	Отклонение T2 от графика (потребитель)
an_Q	303	1.9	Аномалия тепловой мощности

Подобные аномалии не фиксируются прямым сравнением температур с графиком,

но успешно обнаруживаются предложенным методом.

Общая статистика по всем статусам режимов работы представлена в табл. 5.

Анализ результатов показывает, что 56.8% времени параметры ЦТП соответствуют нормативным требованиям. Отклонения T1 (30.1%) указывают на эпизодическое отклонение от температурного графика. Отклонения T2 (13.1%) свидетельствуют о проблемах со стороны потребителей: завышенная температура обратки указывает на неэффективный теплосъем (загрязнение теплообменников), заниженная температура указывает на возможные утечки или некорректную работу погодозависимой автоматики, которая отклоняется от температурного графика. Аномалии Q (1.9%) наиболее часто возникают в переходные периоды и при ремонтных работах, что подтверждает их диагностическую ценность.

Особый интерес представляет анализ периода резкого повышения температуры наружного воздуха (март 2024 г.). Алгоритм зафиксировал рост давления P1 на 8–12% относительно прогнозируемых значений при сохранении нормативных температур T1 и T2. Данное явление объясняется инерционностью системы погодозависимого регулирования, что может вызывать ускоренный износ оборудования. Раннее обнаружение таких режимов позволяет своевременно корректировать настройки и предотвращать аварийные ситуации.

5. ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Точность прогнозирования тепловой мощности, достигнутая полносвязной нейронной сетью ($R^2 = 0.9997$, MAE = 0.0099 Гкал/ч), сопоставима с результатами, полученными для задачи прогнозирования тепловой нагрузки как методами градиентного бустинга в сочетании со статистическими моделями [6], так и более сложными архитектурами глубокого обучения на основе графовых нейронных сетей [7]. Однако прямое численное сопоставление ошибок затруднено: в отличие от [6, 7], где оценка режимов работы строится на суточных или агрегированных по потребителям данных, в настоящей работе используется почасовой ряд по одному объекту с выраженной автокорреляцией и лаговыми признаками, что закономерно способствует более высоким значениям R^2 . Это обстоятельство следует

учитывать при интерпретации результата — высокая точность прогноза является необходимым, но не достаточным условием для содержательных выводов; её практическая ценность подтверждается именно на этапе анализа режимов работы (см. раздел 4.4), а не при сравнении метрик прогноза.

Что касается диагностической составляющей, большинство рассмотренных зарубежных исследований прямо указывают дефицит размеченных данных об авариях как основное ограничение развития методов FDD (fault detection and diagnosis) в теплоснабжении [11, 12]. Часть работ решает эту проблему генерацией синтетических или лабораторно эмулированных размеченных датасетов [14, 15] либо аугментацией данных для дообучения моделей в условиях дефицита разметки [21]. Предложенный в настоящей работе подход решает ту же проблему принципиально иначе — не требует размеченных примеров вовсе, поскольку опирается на сопоставление остатков регрессионной модели с независимым нормативным эталоном (температурным графиком), а не на обучение классификатора неисправностей. Такая опора на регуляторный эталон характерна для систем теплоснабжения постсоветского пространства и не имеет прямых аналогов в проанализированной зарубежной литературе, где подобный нормативный график как отдельный источник эталонных значений, как правило, отсутствует.

Разделение источника отклонения (поставщик тепла — по параметру T1, потребитель — по параметру T2, скрытые нештатные режимы — по Q) также отличает предложенный алгоритм от большинства рассмотренных работ. Например, метод обнаружения аномалий на основе графов ближайших соседей [19] выделяет лишь две категории отклонений — по температуре подачи и по общей эффективности подстанции, — но не разграничивает ответственность между поставщиком и потребителем тепла. Аналогичным образом обзоры [11, 12] показывают, что большинство методов ограничиваются задачей обнаружения (детектированием факта отклонения) без атрибуции источника. В этом смысле предложенная схема ближе к практическим задачам эксплуатации, где разграничение ответственности напрямую определяет дальнейшие действия диспетчера.

Обнаруженная закономерность роста давления в подающем трубопроводе на 8–12% при резком повышении температуры наружного воздуха представляет собой самостоятельный практический результат, не описанный явно в проанализированных источниках; ближайшие по смыслу работы посвящены гидравлической балансировке и управлению переменной скоростью насосов, но рассматривают эту проблему в контексте оптимального управления, а не диагностики предаварийных состояний. Это указывает на то, что связь между инерционностью погодозависимого регулирования и повышенным риском износа сетей — область, слабо освещённая в текущей литературе, и предложенный метод открывает возможность для дальнейшего количественного исследования этого эффекта.

Следует обозначить ряд ограничений, определяющих условия применимости метода на других объектах теплоснабжения.

1. Зависимость от наличия утверждённого температурного графика. Алгоритм анализа T1/T2 использует нормативный график, как эталон. На объектах, где такой график отсутствует, не актуализирован или носит рекомендательный характер, потребуется либо восстановление расчётной зависимости на основе паспортных данных объекта, либо переход на чисто прогнозную схему анализа режима работы (по аналогии с обнаружением аномалий Q).

2. Ограниченный объём и однородность обучающей выборки. Модели обучены на данных за 22 месяца (два отопительных сезона) в одной климатической зоне. Перенос модели на объекты с иной климатической зоной, иной степенью благоустройства зданий, другой схемой подключения (независимая/зависимая) или другим соотношением нагрузок отопления и горячего водоснабжения потребует переобучения моделей и повторной калибровки пороговых значений; прямой перенос обученных весов без адаптации не рекомендуется.

3. Ручной подбор порогов обнаружения аномалий. Пороговое значение 0.05 Гкал/ч для Q выбрано на основе анализа остатков на конкретной валидационной выборке. При масштабировании подхода на большое число разнородных ЦТП статический порог может потребовать замены на адаптивные, самонастраивающиеся схемы, аналогичные подходам, ориентированным на выявление

относительных, а не абсолютных отклонений внутри сети объектов [19, 21].

4. Отсутствие независимой верификации на размеченном наборе неисправностей. Как и в большинстве проанализированных работ [11, 12], в данном исследовании отсутствует эталонный размеченный набор данных аварий для количественной оценки полноты и точности обнаружения (precision/recall). Сопоставление с журналом эксплуатационных событий (раздел 4.4) носит частичный, качественный характер и должно быть дополнено в дальнейших исследованиях количественной оценкой на большем числе подтверждённых инцидентов.

5. Требования к исходным данным. Метод рассчитан на объекты с почасовой (или более частой) регистрацией технологических параметров и метеоданных. На объектах с редкой (суточной) коммерческой фиксацией показаний или неполным составом измеряемых параметров (отсутствие датчиков давления или расхода на одной из сторон) часть диагностических механизмов (в частности, обнаружение аномалий давления) будет недоступна без модификации алгоритма.

Предложенный алгоритм применим, прежде всего, к ЦТП с достаточной историей наблюдений (не менее одного полного отопительного сезона); распространение метода на объекты с иными характеристиками требует дополнительной адаптации, обозначенной выше.

БИБЛИОГРАФИЯ (REFERENCES)

- [1] Bojko E.E., Byk F.L., Myshkina L.S., Ilyushin P.V. Effektivnostteplosnabzheniya municipalnykh obrazovaniy [Heat Efficiency in Municipalities]. *Izvestiya Rossiiskoi akademii nauk. Energetika – Proceedings of the Russian Academy of Sciences. Energy Industry*, 2023, no.4, pp. 66-78. <http://doi.org/10.31857/S0002331023040040>.
- [2] Boyko E., Byk F., Ilyushin P.V., Myshkina L., Filippov S. Approach to modernizing residential-dominated district heating systems to enhance their flexibility, energy efficiency, and environmental friendliness. *Applied Sciences*, 2023, vol. 13, iss. 22, art. 12133. <http://doi.org/10.3390/app132212133>.
- [3] Shkarovskiy A., Kolienko A., Turchenko V. Increasing the efficiency of heat load control in centralized heating networks. *Architecture and Engineering*, 2021, vol. 6, iss. 3, pp. 29–41. <http://doi.org/10.23968/2500-0055-2021-6-3-29-41>

- [4] Golla A., Geis J., Loy T. et al. An operational strategy for district heating networks: application of data-driven heat load forecasts. *Energy Inform*, 2020, vol. 3, art. 22 (2020). <https://doi.org/10.1186/s42162-020-00125-5>
- [5] Chen Z., O'Neill Z., Wen J. et al. A review of data-driven fault detection and diagnostics for building HVAC systems. *Applied Energy*, 2023, vol. 339, art.121030. <https://doi.org/10.1016/j.apenergy.2023.121030>.
- [6] Zhang L., Wen J., Li Y., Chen J., Ye Y., Fu Y., Livingood W. A review of machine learning in building load prediction. *Applied Energy*, 2021, vol. 285, art. 116452. <http://doi.org/10.1016/j.apenergy.2021.116452>.
- [7] Xue P., Jiang Y., Zhou Z., Chen X., Fang X., Liu J. Multi step ahead forecasting of heat load in district heating systems using machine learning algorithms. *Energy*, 2019, vol. 188, art. 116085. <http://doi.org/10.1016/j.energy.2019.116085>.
- [8] Idowu S., Saguna S., Åhlund C., Schelén O. Applied machine learning: Forecasting heat load in district heating system. *Energy and Buildings*, 2016, vol. 133, pp. 478–488. <http://doi.org/10.1016/j.enbuild.2016.09.068>.
- [9] Zhu J., Yang J., Peng M., Liang X., Ben C. Deep learning-driven heat load prediction: Investigating the impacts of optimizer and learning rate scheduler strategies. *Energy*, 2025, vol. 340, art. 139195. <http://doi.org/10.1016/j.energy.2025.139195>.
- [10] Ding H., Zhan Y. Hourly Heat Load Prediction for District Heating Systems based on Long Short-Term Memory Network. *International Core Journal of Engineering*, 2026, vol. 12, iss. 4, pp. 529-535. [https://doi.org/10.6919/ICJE.202604_12\(4\).0055](https://doi.org/10.6919/ICJE.202604_12(4).0055).
- [11] Ahmad T., Chen H., Huang R., Guo Y., Wang J., Shair J., Akram H.M.A., Mohsan S.A.H., Kazim M. Supervised based machine learning models for short, medium and long-term energy prediction in distinct building environment. *Energy*, 2018, vol. 158, pp. 17–32. <http://doi.org/10.1016/j.energy.2018.05.169>.
- [12] Neumayer M., Stecher D., Grimm S., Maier A., Bucker D., Schmidt J. Fault and anomaly detection in district heating substations: A survey on methodology and data sets. *Energy*, 2023, vol. 276, art. 127569. <http://doi.org/10.1016/j.energy.2023.127569>.
- [13] Shahid Z.K., Saguna S., Åhlund C. Autoencoders for Anomaly Detection in Electricity and District Heating Consumption: A Case Study in School Buildings in Sweden. Proc. Int. Conf. “2023 IEEE International Conference on Environment and Electrical Engineering and 2023 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe)”, Madrid, Spain, 2023, pp. 1–8. <http://doi.org/10.1109/EEEIC/ICPSEurope57605.2023.10194605>.
- [14] Mbiydzanyuy G., Sundell H. Pattern Detection in Abnormal District Heating Data. *Machine Learning, Optimization, and Data Science: 10th International Conference*, LOD 2024, Revised Selected Papers, Part I (Lecture Notes in Computer Science). Springer, 2024, pp. 224–239. http://doi.org/10.1007/978-3-031-82481-4_16.
- [15] Fatehijananloo M., McArthur J. Explainable Machine Learning for Hospital Heating Plants: Feature-Driven Modeling and Analysis. *Buildings*, 2026, vol. 16, art. 397. <http://doi.org/10.3390/buildings16020397>.
- [16] Chen Z., Xiao F., Guo F., Yan J. Interpretable machine learning for building energy management: A state-of-the-art review. *Advances in Applied Energy*, 2023, vol. 9, art. 100123. <http://doi.org/10.1016/j.adapen.2023.100123>.
- [17] Park S., Moon J., Hwang E. Explainable Anomaly Detection for District Heating Based on Shapley Additive Explanations. Proc. Int. Conf. “2020 IEEE International Conference on Data Mining Workshops (ICDMW)”, Sorrento, Italy, 2020, pp. 762–765. <http://doi.org/10.1109/ICDMW51313.2020.00111>
- [18] Dreven J., Cheddad A., Ghazi A.N., Alawadi S., Al Koussa J., Vanhoudt D. HEAT: Hierarchical-constrained Encoder-Assisted Time series clustering for fault detection in district heating substations. *Energy and AI*, 2025, vol. 21, art. 100548. <http://doi.org/10.1016/j.egyai.2025.100548>.
- [19] Li Z., Luo Z., Zhang N., Lin X., Huang W., Feng E., Zhong W. Investigation of Predictive Regulation Strategy of Secondary Loop in District Heating Systems. *Sustainability*, 2023, vol. 15, art. 3524. <http://doi.org/10.3390/su15043524>.
- [20] Che Z., Sun J., Na H., Yuan Y., Qiu Z., Du T. A novel method for intelligent heating: On-demand optimized regulation of hydraulic balance for secondary networks. *Energy*, 2023, vol. 282, art. 128900. <http://doi.org/10.1016/j.energy.2023.128900>.
- [21] Chicherin, S. Hydraulic Balancing of District Heating Systems and Improving Thermal Comfort in Buildings. *Energies*, 2025, vol. 18, iss. 5, art. 1259. <https://doi.org/10.3390/en18051259>.
- [22] Mirnaghi M.S., Haghighat F. Fault detection and diagnosis of large-scale HVAC systems in buildings using data-driven methods: A comprehensive review. *Energy and Buildings*, 2020, vol. 229, art. 110492. <http://doi.org/10.1016/j.enbuild.2020.110492>.
- [23] Singh V., Mathur J., Bhatia A. A comprehensive review: Fault detection, diagnostics, prognostics, and fault modeling in HVAC

- systems. *International Journal of Refrigeration*, 2022, vol. 144, pp. 283–295. <http://doi.org/10.1016/j.ijrefrig.2022.08.017>
- [24] Wang Z., Hong T., Piette M.A. Building thermal load prediction through shallow machine learning and deep learning. *Applied Energy*, 2020, vol. 263, art. 114683. <http://doi.org/10.1016/j.apenergy.2020.114683>.
- [25] Galicia A., Talavera Llamas R., Troncoso A., Koprinska I., Martínez Álvarez F. Multi step forecasting for big data time series based on ensemble learning. *Knowledge-Based Systems*, 2019, vol. 163, pp. 830–841. <https://doi.org/10.1016/j.knosys.2018.10.009>.
- [26] Gu J., Wang J., Qi C., Min C., Sundén B. Medium-term heat load prediction for an existing residential building based on a wireless on-off control system. *Energy*, 2018, vol. 152, pp. 709–718. <http://doi.org/10.1016/j.energy.2018.03.179>.
- [27] Zhao A., Dong F., Xue X., Xi J., Wei Y. Optimal control for hydraulic balance of secondary network in district heating system under distributed architecture. *Energy and Buildings*, 2023, vol. 290, art. 113030. <http://doi.org/10.1016/j.enbuild.2023.113030>.
- [28] Dang L. M., Nguyen L. Q., Nam J., Nguyen T. N., Lee S., Song H.-K., Moon H. Fifth generation district heating and cooling: A comprehensive survey. *Energy Reports*, 2024, vol. 11, pp. 1723–1741. <http://doi.org/10.1016/j.egyr.2024.01.03>.
- [29] Jiana Y., Lia Y., Liua Z., Lib R. Real operation pattern of district heating system and its heating effects. *Procedia Engineering*, 2015, vol. 121, pp. 1741–1748. <http://doi.org/10.1016/j.proeng.2015.09.146>.
- [30] Chambers J., Narula K., Sulzer M., Patel M. K. Mapping district heating potential under evolving thermal demand scenarios and technologies: A case study for Switzerland. *Energy*, 2019, vol. 176, pp. 682–692. <http://doi.org/10.1016/j.energy.2019.04.04>.

Сведения об авторах.



Дворцевой Александр Игоревич, Новосибирский государственный технический университет, к.т.н., доцент. Область научных интересов: анализ технологических данных и диагностика оборудования
ORCID: 0009-0004-3817-8813
E-mail: dvorcevoj@corp.nstu.ru



Яковина Ирина Николаевна, Новосибирский государственный технический университет, к.т.н., доцент. Область научных интересов: исследование эффективности применения методов машинного обучения.
ORCID: 0000-0002-3265-8865
E-mail: yakovina@corp.nstu.ru



Боруш Олеся Владимировна, Новосибирский государственный технический университет, д.т.н., профессор. Область научных интересов: новые схемные решения для энергетических установок.
ORCID: 0000-0001-5080-0872
E-mail: olesyaborush@yandex.ru



Бойко Екатерина Евгеньевна, Новосибирский государственный технический университет, к.т.н., доцент. Область научных интересов: системы управления теплоснабжением; мультэнергетические системы теплоснабжения.
ORCID: 0000-0001-6142-039X
E-mail: e.boyko1991@yandex.ru



Яньков Павел Александрович, Новосибирский государственный технический университет, бакалавр. Область научных интересов: прикладное применение методов машинного обучения в промышленности.
ORCID: 0009-0006-6042-6902
E-mail: yankov.2022@stud.nstu.ru



Виндемут Данил Юрьевич, Новосибирский государственный технический университет, магистрант. Область научных интересов: Энергосбережение и оптимизация систем теплоснабжения.
ORCID: 0009-0009-1032-6024
E-mail: danil-golikov-2004@mail.ru