

**Data-Driven Modeling and Analysis of the Highly Dynamic  
User-Level Behavior in Heat Networks  
to Enhance the Efficiency and Flexibility**

Doctoral Thesis (Dissertation)  
to be awarded the degree  
Doctor of Engineering (Dr.-Ing.)

submitted by  
M.Eng. Samanta Alena Weber

approved by  
the Faculty of Engineering and Economics  
at Clausthal University of Technology,

Date of oral examination  
30th June 2026

Dean

Prof. Dr.-Ing. Armin Lohrengel

Chairperson of the Board of Examiners

Prof. Dr.-Ing. habil. Philip Jaeger

Supervising tutors/Reviewers

Prof. Dr. mont. Dr. rer. nat. Michael Fischlschweiger

Prof. Dr.-Ing. Dirk Volta

Reviewer

Prof. Dr.-Ing. Ines Hauer

## Acknowledgements

First, I would like to express my sincere thanks to my supervisors. Prof. Dr. mont. Dr. rer. nat. Michael Fischlschweiger was inspiring throughout the work on my cumulative dissertation and my academic journey. I further specially thank him for the constructive discussions. Likewise, I want to express my deep gratitude to Prof. Dr.-Ing. Dirk Volta, who supported me with his time and provided the resources to implement my research ideas.

Second, I thank Prof. Dr.-Ing. Ines Hauer for her efforts in assessing this work. Furthermore, I want to thank Prof. Dr.-Ing. Jens Geisler and Ulf Rieck-Blankenburg, each for their contributions to one of the publications related to this cumulative dissertation.

Third, I'd like to recognize the employees of Clausthal University of Technology. I would like to thank the Promotionskommission for their valuable work. Moreover, I'd like to mention Nicola Henze for facilitating my remote work. Likewise, I'm grateful to my colleagues and friends at the University of Applied Sciences Flensburg, especially Mika Suwe, for his friendship.

Fourth, I would like to thank the Stadtwerke Flensburg GmbH. I'd like to especially mention Thomas Räther, Dirk Roschek, Maike Kristensen, Vadim Korogodski, Varina Oesterle, Klaus Reinwand, Dennis Faulhaber, Hella Plechinger, Thomas Heß, Artur Scheiermann, and Manuel Schröder.

I further want to acknowledge the Society of Energy and Climate Protection in Schleswig-Holstein (EKSH) for funding the project "Szenarienentwicklung für die emissionsoptimierte Energieversorgung einer Fernwärmestadt (CERO2)" Nr. 8/12-57, and for the scholarship Nr. 14/12-34, which helped finance this work. I also thank Dr. Thies Rasmus Popp for his encouragement that led me to discover my abilities in Science Slamming.

Moreover, I appreciate the work of the editors and anonymous reviewers of the three publications that are part of this cumulative dissertation. Furthermore, I would like to acknowledge the discussions I had with Dr. Gesina Schwalbe, which broadened my understanding of Machine Learning. Moreover, I thank Frank Karolius for the always interesting conversation.

Finally, I would like to thank my parents, my friends, and, above all, my beloved, wonderful, and loving husband. You have had my back. Thank you!



**Hochschule  
Flensburg**  
University of  
Applied Sciences



Gesellschaft für  
Energie und Klimaschutz  
Schleswig-Holstein

## Abstract

The growing urgency to address climate change and resource scarcity has accelerated efforts toward the energy transition. In this context, heat networks are gaining importance, as they are considered one of the most efficient supply solutions for areas with high heat demand density. However, to enable the transition to climate-neutral operation, a transformation phase is required in which existing heat networks are improved in efficiency and flexibility.

Against this background, the scientific literature proposes closing the existing information gap between operators and users to enhance heat network control strategies. However, making user-level information accessible poses various challenges. Initial successes have been achieved in modeling the aggregated heat demand of large residential complexes. In contrast, the modeling of highly dynamic user behavior, as observed in buildings with few regular occupants, such as single-family homes, has so far been investigated only to a limited extent. This cumulative dissertation pursues the overarching objective of developing suitable modeling and analysis approaches that make information on highly dynamic user-level behavior accessible. This aims to enable the use of this information to enhance efficiency and flexibility in heat networks.

Methodologically, the necessary information to achieve this overarching objective is developed step by step in a systematic manner. To this end, an automated approach based on thermodynamic principles is developed to filter faulty and non-representative datasets. This improves data quality by generating cleaned datasets. Furthermore, relevant and non-redundant input parameters for modeling highly dynamic user behavior are identified. Subsequently, the results are integrated into a Long Short-Term Memory modeling framework.

The results achieved with the developed methodology first enable the automated identification of faulty substations at the user level. For practical application, the method also quantifies the negative impact of the detected faults on the system. The insights gained improve data quality and contribute to enhanced fault management at the user level in heat networks, thereby enabling the increase of efficiency and flexibility. Furthermore, the findings demonstrate how data-driven modeling approaches can be applied in heat networks. This enables accessing the necessary information for improved understanding of highly dynamic user behavior in heat networks. Model optimizations and a systematic error analysis are carried out.

These findings on highly dynamic user-level behavior are relevant to modeling heat networks and to integrating this type of information into practical network operation, particularly for improving control strategies.

Moreover, this cumulative dissertation identifies starting points for future research. For example, continuous model updating and self-learning modeling approaches could be incorporated into heat network control strategies to adapt to changing conditions. The transferability of the modeling strategy developed in this cumulative dissertation to other model regions and heat networks also represents a relevant question for future research.

## Zusammenfassung

Die zunehmende Dringlichkeit, auf den Klimawandel sowie auf die Ressourcenknappheit zu reagieren, hat die Bemühungen um die Energiewende beschleunigt. In diesem Rahmen gewinnen Wärmenetze an Bedeutung, da sie als eine der effizientesten Versorgungslösungen für Gebiete mit hoher Wärmenachfragedichte gelten. Um jedoch den Übergang zu einer klimaneutralen Betriebsweise zu ermöglichen, ist eine Transformationsphase erforderlich, in der bestehende Wärmenetze hinsichtlich ihrer Effizienz und Flexibilität verbessert werden.

Vor diesem Hintergrund wird in der wissenschaftlichen Literatur vorgeschlagen, die bestehende Informationslücke zwischen Betreibern und Nutzern zu schließen, um die Regelungsstrategie von Wärmenetzen zu verbessern. Das Zugänglichmachen nutzerseitiger Informationen ist jedoch mit verschiedenen Herausforderungen verbunden. Erste Erfolge hierzu wurden bei der Modellierung des aggregierten Wärmebedarfs größerer Wohnanlagen erzielt. Im Gegensatz dazu wurde die Modellierung des hochdynamischen Verhaltens, das sich bei Gebäuden mit wenigen regelmäßigen Nutzern wie etwa Einfamilienhäusern zeigt, bislang kaum systematisch untersucht. Die vorliegende kumulative Dissertation setzt das übergeordnete Ziel, geeignete Modellierungs- und Analyseansätze zu entwickeln, die Informationen zum hochdynamischen Verhalten der Nutzerebene zugänglich machen, um diese zur Steigerung der Effizienz und Flexibilität in Wärmenetzen nutzen zu können.

Methodisch werden zu Beginn die notwendigen Informationen zur Erreichung des übergeordneten Ziels schrittweise und systematisch erarbeitet. Dazu wird ein automatisierter und auf thermodynamischen Grundprinzipien fundierter Ansatz zur Herausfilterung fehlerhafter und nicht repräsentativer Datensätze entwickelt, um die Datenqualität zu verbessern, indem bereinigte Datensätze erzeugt werden. Weiterhin werden relevante, nicht redundante Eingangsparameter zur Modellierung des hochdynamischen Nutzerverhaltens bestimmt. Daraufhin werden die Ergebnisse in ein Long-Short-Term-Memory-Modellgerüst integriert.

Die mit der entwickelten Methodik erzielten Ergebnisse ermöglichen zunächst die automatisierte Identifizierung fehlerhafter Wärmeübergabestationen auf der Nutzerebene. Für die praktische Anwendung quantifiziert die Methode zudem die negativen Auswirkungen der detektierten Fehler auf das System. Die gewonnenen Erkenntnisse steigern die Datenqualität und tragen zu einem verbesserten Fehlermanagement auf der Nutzerebene in Wärmenetzen bei, wodurch Effizienz und Flexibilität erhöht werden können. Weiterhin zeigen die gewonnenen Erkenntnisse, wie die Anwendung datengetriebener Modellierungsansätze in Wärmenetzen erfolgen kann. Damit können die notwendigen Informationen für ein verbessertes Verständnis des hochdynamischen Nutzerverhaltens in Wärmenetzen generiert werden. Es werden Modelloptimierungen und eine systematische Fehleranalyse durchgeführt.

Diese Ergebnisse zum hochdynamischen Verhalten der Nutzerebene sind für die Modellierung von Wärmenetzen sowie für die Integration dieser Art von Informationen in den praktischen Netzbetrieb relevant, insbesondere zur Verbesserung der Regelungsstrategie.

Weiterhin ergeben sich in dieser kumulativen Dissertation Ansatzpunkte für zukünftige Forschungsfragen. Beispielsweise können die fortlaufende Aktualisierung und die selbstlernende Modellierung in die Regelungsstrategien von Wärmenetzen einfließen, um sich an sich verändernde Bedingungen anzupassen. Auch die Übertragbarkeit der in dieser kumulativen Dissertation entwickelten Modellierungsstrategie auf andere Modellregionen und Wärmenetze im Vergleich ist eine für die Zukunft relevante Fragestellung.

**Table of Contents**

|                                                                                                                                                                      |             |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| <b>Acknowledgements.....</b>                                                                                                                                         | <b>III</b>  |
| <b>Abstract .....</b>                                                                                                                                                | <b>V</b>    |
| <b>Zusammenfassung.....</b>                                                                                                                                          | <b>VI</b>   |
| <b>Table of Contents .....</b>                                                                                                                                       | <b>VII</b>  |
| <b>List of Figures .....</b>                                                                                                                                         | <b>VIII</b> |
| <b>1 Introduction.....</b>                                                                                                                                           | <b>1</b>    |
| <b>2 Objective and Subobjectives of this Cumulative Dissertation .....</b>                                                                                           | <b>6</b>    |
| 2.1 Overarching Objective of this Cumulative Dissertation.....                                                                                                       | 6           |
| 2.2 Subobjectives of this Cumulative Dissertation .....                                                                                                              | 6           |
| 2.2.1 Subobjective 1: Clustering- and Statistic-Based Approach for Detection and Impact Evaluation of Faults in End-User Substations of Thermal Energy Systems ..... | 6           |
| 2.2.2 Subobjective 2: Feature Selection for Specific Prediction Targets at the User Level in a District Heating Network .....                                        | 7           |
| 2.2.3 Subobjective 3: Dynamic User-Level Parameter Modeling in Heat Networks with Deep Learning Long Short-Term Memory .....                                         | 8           |
| <b>3 Overview of Publications.....</b>                                                                                                                               | <b>9</b>    |
| 3.1 Publication 1: Clustering- and Statistic-Based Approach for Detection and Impact Evaluation of Faults in End-User Substations of Thermal Energy Systems ....     | 9           |
| 3.1.1 Author Contribution Statement of Publication 1.....                                                                                                            | 9           |
| 3.1.2 Publication 1.....                                                                                                                                             | 9           |
| 3.1.3 Supplementary Material for Publication 1 .....                                                                                                                 | 27          |
| 3.2 Publication 2: Feature Selection for Specific Prediction Targets at the User Level in a District Heating Network .....                                           | 32          |
| 3.2.1 Author Contribution Statement of Publication 2.....                                                                                                            | 32          |
| 3.2.2 Publication 2.....                                                                                                                                             | 32          |
| 3.2.3 Supplemental Material of Publication 2.....                                                                                                                    | 50          |
| 3.3 Publication 3: Dynamic User-Level Parameter Modeling in Heat Networks with Deep Learning Long Short-Term Memory .....                                            | 55          |
| 3.3.1 Author Contribution Statement of Publication 3.....                                                                                                            | 55          |
| 3.3.2 Publication 3.....                                                                                                                                             | 56          |
| 3.3.3 Supplemental Material of Publication 3.....                                                                                                                    | 69          |
| <b>4 Discussion, Conclusions, and Outlook .....</b>                                                                                                                  | <b>80</b>   |
| <b>5 References.....</b>                                                                                                                                             | <b>84</b>   |

**List of Figures**

|                                                                                                |    |
|------------------------------------------------------------------------------------------------|----|
| FIGURE 1 SCIENCEDIRECT SEARCH RESULTS FOR DIFFERENT KEYWORDS ON PIPE-BOUND HEAT SUPPLY .....   | 2  |
| FIGURE 2 SCIENCEDIRECT SEARCH RESULTS FOR KEYWORDS ON DATA-DRIVEN PROCESSING STRATEGIES .....  | 3  |
| FIGURE 3 STRUCTURE OF THIS RESEARCH WITH SUBOBJECTIVES, DATA, METHODS, AND CONTRIBUTIONS ..... | 82 |

# 1 Introduction

With the growing need to mitigate the effects of climate change and address resource scarcity, efforts to achieve the energy transition have increased significantly in recent years. In this context, heat networks have come into focus, as they are commonly considered one of the most efficient options for heat supply in areas of high demand density [12, 38, 41]. Yet, while the European Union covered 47.5% of its electric energy demand with renewables in 2024, only 26.7% of its thermal energy demand was met by CO<sub>2</sub>-neutral sources [23]. Given the anticipated high efficiency, the strategic importance of heat networks could increase significantly in future energy systems. Figure 1 substantiates the growing interest in the relevance of heat networks. It illustrates the number of search hits on ScienceDirect<sup>1</sup> [62] by publication year for the selected keywords “district heating”, “thermal energy system”, “heat network”, “heat system”, and “heat grid” as of 29<sup>th</sup> January 2026. Moreover, in Figure 1 a), relevant geopolitical events and their date of occurrence are highlighted. As the term “district heating” stands out significantly in Figure 1 a), Figure 1 b) shows a magnification of the results when excluding this term, enabling a distinction for the remaining keywords.

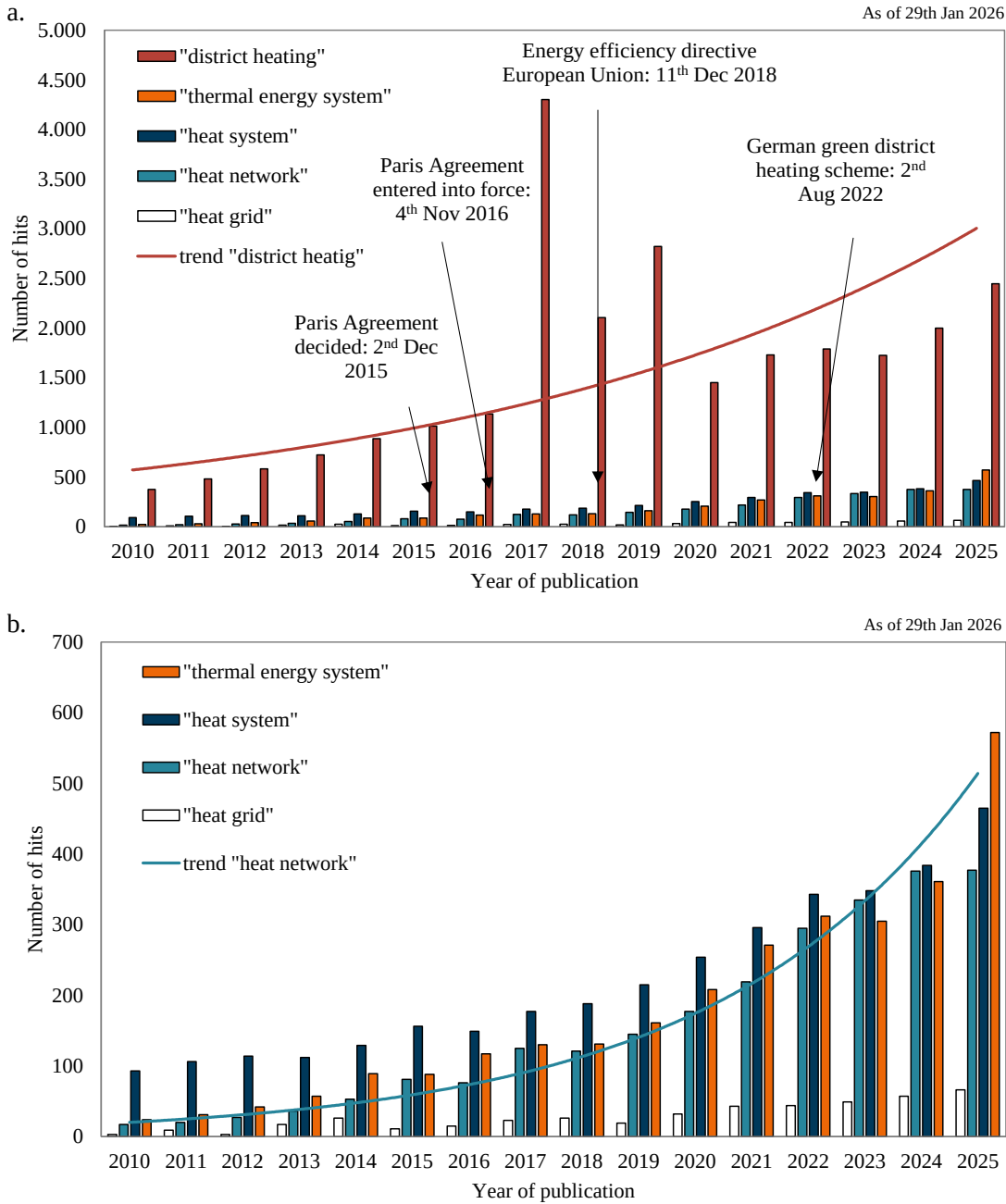
As indicated in Figure 1 a), (e.g., political) incentives lead to an increase in research, which enables a linking of the point in time to possible reasons for changes in the number of publications that exist. Figure 1 a) suggests that interest in the subject of district heating increased significantly in 2017. After this year, the number of publications first decreased but has risen again since 2020. In comparison, Figure 1 b) shows that scientific interest in pipe-bound heat supply, as reflected in the remaining keywords, has continuously increased. Generally, the investigation confirms the growing relevance of the topic in research. Regarding Figure 1, the specific interest in heat networks appears to be driven by their potential contribution to the heat transition. Yet efforts to integrate renewables pose particular challenges for existing heat networks. Such heat networks need to transform to meet future requirements and achieve lower operational temperatures and greater flexibility, as outlined in the concept of the 4th generation of district heating [45].

As operations are intended to be optimized to meet future requirements, the limitations of conventional operational practices become increasingly apparent. Traditionally, heat network operation was a knowledge-based, reactive task primarily aimed at balancing heat demand in response to weather conditions [74]. This can enable sufficient control strategies but not optimized operations [74]. Improved operational strategies for heat networks were found to reduce heat losses by up to 6-7 %, not yet including the efficiency gains of the supply plant [14]. Thus, it is emphasized that traditional approaches can fail to account for all influencing factors and, therefore, exhibit insufficient performance [31]. As a specific shortcoming of traditional operational concepts, Ref. [61] notes that these principles are often feed-forward, meaning they do not allow back-coupling of user information.

In response to these limitations, one specific suggestion to further enhance the efficiency of heat networks is to close the information gap between operators and users by establishing bottom-up control [51, 80]. For example, accurate predictions of substations' supply-side parameters could reduce existing safety margins without risking undersupply [51]. This underscores the critical need to make user-level heat network information accessible. The information obtained should be valuable for enhancing the control strategy by providing insights into the efficiency and flexibility. Hence, there is a current need for reliable, high-quality information on user-level behavior in heat networks.

---

<sup>1</sup> Particular emphasis is placed on ScienceDirect, as the literature review suggests that, to the best of the author's knowledge, this platform has published the largest number of relevant articles to date.



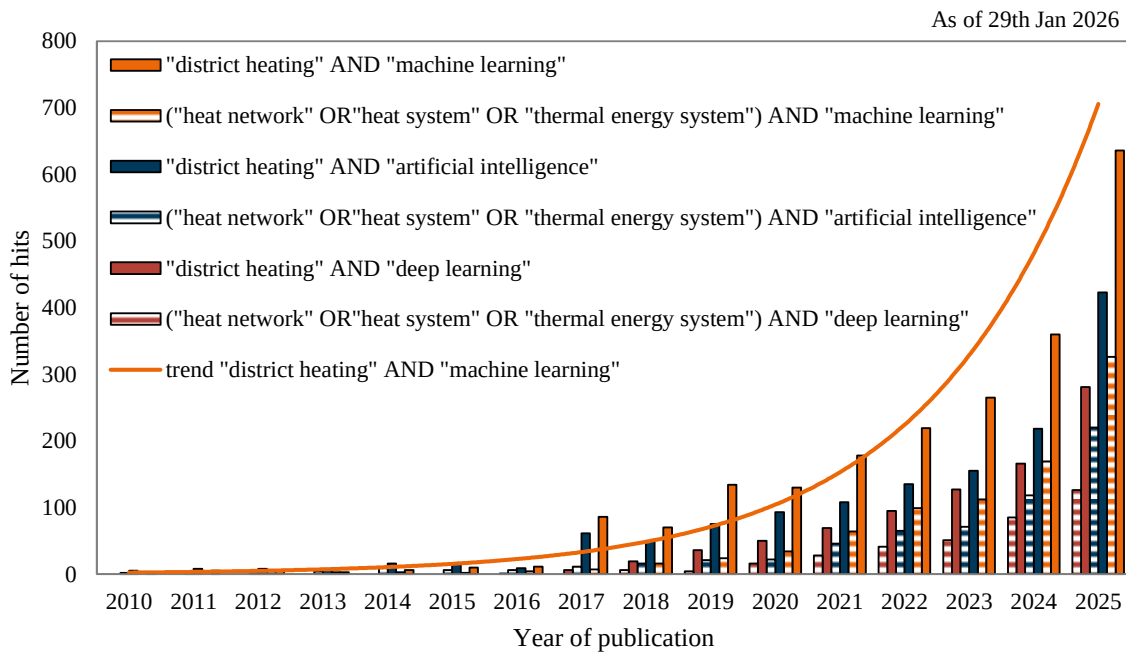
**Figure 1** ScienceDirect search results for different keywords on pipe-bound heat supply for a) “district heating”, “thermal energy system”, “heat network”, “heat system”, and “heat grid” with occurrence dates of relevant geopolitical events, and trend for “district heating” b) magnification excluding term “district heating” and including a trend for “heat network”

One option to make user-level information in heat networks accessible for practical integration into control concepts is to develop appropriate modeling and analysis methods for user-level parameters. Primary existing successful applications provide valuable motivation for further investigation. In particular, aggregated heat demand for residential complexes has already been successfully modeled using Machine Learning (ML) frameworks [58]. However, modeling individual demand is considerably more challenging than modeling aggregated demand from a larger number of regular users, owing to the increased influence of individual occupants [58]. Hence, modeling buildings with few regular occupants (e.g., single-family homes) remains a challenge [32, 35], because the lower the number of regular occupants of a building is, the higher the modeling error becomes [32]. Therefore,

for heat demand data in particular, certain studies exclude detached houses [9, 22]. Thus, knowledge of modeling and analysis for this particular type of user-level behavior in heat networks is limited [5] and requires further investigation.

To target the highly dynamic user-level behavior in heat networks, the modeling and analysis methodology must be appropriately selected. For instance, existing approaches use physical models to estimate parameters, which are specific to the modeled heat network structure and are challenging to set up [51]. Given the shortcomings of existing strategies [74], data-driven approaches, such as ML, have gained significance in comparison [31]. This is because they provide a systematic approach to harnessing the potential of data by using algorithms that process large-scale datasets and yield interpretable conclusions, but require limited detail on infrastructure (such as network, buildings, and equipment) [54]. Furthermore, ML concepts make decisions based on historical (real) data and are highly effective in modeling nonlinear behavior [54]. Thus, ML is considered a method that will “revolutionize” district heating through increased automation and greater flexibility [54]. The advantages of data-driven modeling support its application to heat networks: ML models enable a high degree of automation in model building [51], thereby simplifying adaptation through retraining [54].

To substantiate the increase in interest arising from the potential benefits of applying ML to heat network data, Figure 2 presents results for the combination of keywords related to heat supply with “artificial intelligence”, “machine learning”, and “deep learning”. Based on the results in Figure 1, the terms “heat network” and “heat system” are combined, while the seemingly less significant term “heat grid” is excluded.



**Figure 2 ScienceDirect search results for keywords on data-driven processing strategies and pipe-bound heat supply for “district heating”, “heat network” OR “heat system” OR “thermal energy system” AND “machine learning”, “artificial intelligence” and “deep learning”, and trend for “district heating” AND “machine learning”**

Figure 2 indicates the highest number of hits for the term “machine learning” in recent years, but “artificial intelligence” is a standard reference as well. Furthermore, Figure 2 showcases that “deep learning” has gained importance since 2017. As the upward trend underscores, interest in combining ML and heat network data has been increasing.

However, Figure 2 also visualizes that due to limited data availability [5], heat networks have only recently become a focus of data-driven investigations [26]. In existing work on data-driven concepts applied in heat networks, a focus on modeling the heat network inputs is discernible, especially regarding the heat load prediction (e.g., Ref. [9, 12, 15, 16, 20, 26, 34, 41, 44, 46, 56, 65, 68, 70]). For this modeling target, data have been available for a considerable period. In contrast, user-level heat network data have only recently become available, despite their suggested high relevance for enhancing heat network efficiency and flexibility. Specifically in Germany, the “Fernwärme- oder Fernkälte-Verbrauchserfassungs- und -abrechnungsverordnung” (FFVAV) has required the use of remotely accessible data-collection units at the user level for all newly installed facilities since 2021 [8]. Retrofitting of existing facilities is required by 2027 [8]. Yet, even though user-level data are becoming available, the specific information content captured in the data that will enable efficiency and flexibility gains in the heat network has yet to be obtained, e.g., by applying ML. Because data availability is a significant barrier to data-driven modeling [5], research on combining ML methods with user-level data from heat networks is scarce [51].

Hence, although ML offers the aforementioned advantages, applying this methodological framework to user-level data that reflects highly dynamic behavior still requires specific knowledge. The uniqueness of experimental data necessitates individualized information about the most sophisticated data-driven model-building process. It is uncommon for a single optimal approach to exist, as the choice of ML architecture and (hyper-)parameterization is typically domain- and dataset-specific for a given modeling objective [29]. Therefore, modeling newly suggested or scarcely investigated parameters, such as those that reflect the highly dynamic user-level behavior of heat networks, requires intensive inquiry into the conceptualization of the modeling framework.

In this context, selecting an appropriate modeling architecture is particularly critical [54]. Regarding the choice of ML modeling architecture, it is generally stated that the essential capability of data-driven approaches nowadays should be a “memory” [74]. This is achieved by adding recurrent connections, which is why Recurrent Neural Networks (RNNs) have gained popularity [74]. However, the effectiveness of the traditional concept is considered limited [74]. Thus, Ref. [12] identifies the currently preferred and most successful algorithmic structures within the ML spectrum for the specific application of heat load forecasting, namely Deep Neural Networks (DNNs) in general and the Long Short-Term Memory (LSTM) architecture in particular. The LSTM uses not just single recurrent connections as input for the neurons, as traditional RNNs do, but blocks that contain one or more memory cells [74]. These memory cells are connected in a recurrent architecture and comprise input, output, and forget units, which can be used to write, read, and reset the cells, respectively [74]. Despite the success of the LSTM approach for heat load modeling [12], it has not yet been holistically tested in combination with data reflecting the highly dynamic user-level behavior of heat networks.

Likewise, the data must be selected in a unique combination with the ML architecture. For instance, this involves selecting inputs [15, 16, 46, 56, 68, 70], data pre-preparation [16, 20, 28, 31, 47, 55, 63, 76], and optimizing hyperparameters [43, 64]. In general, for ML modeling, clean, sufficiently sized [5] datasets are required. Therefore, data scarcity remains a significant barrier [5]. This includes datasets that exhibit fault-induced patterns, sudden changes in behavior, and large data gaps. Therefore, anomalous recordings pose challenges, as models trained on such datasets can be misleading. Hence, automatically detecting and filtering outlier datasets is essential for improving data quality in ML modeling. Beyond predictive and modeling capabilities, ML enables such deeper analytical insights. Herein, a common approach is to cluster users by their load profiles [25, 53]. This enables the identification of outliers [9] and the subsequent automated filtering of datasets unsuitable for modeling.

In summary, a specific research gap is the need to make information on the highly dynamic user-level behavior in heat networks accessible to inform enhanced control strategies that increase efficiency and flexibility. Addressing this gap, however, entails substantial methodological challenges. To access this type of user-level information, modeling and data analysis are key methodological instruments. In particular, ML appears to be a promising methodological framework. However, applying a data-driven approach requires further investigation, specifically for the highly dynamic behavior captured in user-level data. The following presents the research objective arising from the identified research gap, which informs the development and systematic implementation of the research methodology outlined in this cumulative dissertation.

## 2 Objective and Subobjectives of this Cumulative Dissertation

Overall, this cumulative dissertation focuses on the so-far scarcely investigated user-level data from primarily single-family homes that capture high dynamics driven by individual demand and behavioral patterns. Making these accessible could improve the control strategy for heat networks by closing the information loop between operators and users. This is intended to reduce safety margins without risking undersupply, thereby enhancing the efficiency and flexibility of the heat network.

### 2.1 Overarching Objective of this Cumulative Dissertation

Below, the overarching research objective, which is derived from the identified research gap, is formulated.

---

This cumulative dissertation pursues the overarching objective of developing suitable modeling and analysis approaches to make information on highly dynamic user-level behavior of heat networks accessible, thereby enabling the enhancement of efficiency and flexibility.

---

Inquiring into the overarching research objective requires methodological implementation. The methodology in this cumulative dissertation is developed in an interdependent relationship with the available data. In the investigated model region in northern Germany, data are recorded at the building level at approximately hourly intervals for supply and return temperatures and volume flow at the primary side. These recordings can potentially provide demand information and quantifiable feedback on supply efficiency and flexibility. Hence, they serve as one means of closing the control loop and bridging the information gap between users and the operator of the heat network. Therefore, these parameters are selected as the modeling targets and center of analysis for this cumulative dissertation. Because the building stock of the model region comprises a significant share of single-family homes and other residential buildings with few regular occupants, the data reflect highly dynamic behavioral patterns, as expected for these building types. Thus, modeling and analyzing the selected target parameters are consistent with the overarching objective of making information about highly dynamic user-level behavior accessible.

To pursue the overarching research objective, a research design with three subobjectives was developed. These three subobjectives align with three contributions to the literature associated with this cumulative dissertation, which aim to make information on highly dynamic user-level behavior in heat networks accessible to achieve efficiency and flexibility gains.

### 2.2 Subobjectives of this Cumulative Dissertation

In the following, the individual subobjectives are explained and placed in the context of the overarching objective of this cumulative dissertation.

#### 2.2.1 Subobjective 1: Clustering- and Statistic-Based Approach for Detection and Impact Evaluation of Faults in End-User Substations of Thermal Energy Systems

In data-driven modeling, high-quality data is essential for model performance. To model user-level behavior, the data should be representative of the expected user patterns. Hence, the first investigation in this cumulative dissertation develops a method to automatically filter datasets unsuitable for modeling due to anomalies in the recorded data, such as fault-induced behavior. To pursue this aim and, in parallel, provide a practical benefit, subobjective 1 is formulated to increase automation in fault management of user-level substations in heat networks. Effective fault management is an essential step toward achieving the required transformation for heat defossilization. Yet existing fault

detection processes exhibit shortcomings that strain operators' resources for fault management. Hence, increasing automation in this context is intended to advance fault management at the building level, thereby enhancing heat network efficiency and flexibility.

As noted, to avoid overstraining the operator's limited resources for fault management [6], certain practical constraints should receive more consideration [60, 79]. The existing literature identifies, for example, reducing manual processing steps [6], enabling the prioritization of fault-elimination measures [49], using only accessible (i.e., primary-side) data [60], and taking into account reliability and cost-effectiveness, as well as transferability and scalability across different heat networks [79]. To address this, this cumulative dissertation considers the aforementioned practical requirements for the feasibility of fault management. Additionally, the methodological implementation for detecting anomalous patterns should enable the automated identification and filtering of datasets unsuitable for data-driven model building.

In the publication pursuing subobjective 1, it is investigated how automatically detectable indicators for existing faults in building-level substations can be derived from user-level temperature data. Moreover, the negative effect of the fault on the heat network is quantified by combining information on existing faults with the volume-flow data. Publication 1 formulates a fault management approach that aims for practical feasibility and enables automated fault detection in user substations. Furthermore, this step provides the methodological framework for improving data quality in data-driven modeling by automatically removing datasets unsuitable for this purpose.

## **2.2.2 Subobjective 2: Feature Selection for Specific Prediction Targets at the User Level in a District Heating Network**

To develop any modeling framework, it is necessary to identify inputs that effectively yield the desired outputs. The inputs of an ML model are referred to as “features”. Transforming raw data of potential inputs into datasets processible by an ML algorithm is called “feature engineering.” Determining a subset of features that are non-redundant and of as high relevance as possible for the target of prediction is the so-called “feature selection.”

The choice of predictors decisively influences model performance [56]. However, research on the most sophisticated selection of inputs tends to be limited for heat network modeling in general and for highly dynamic user-level behavior in particular [15, 46]. Hence, the second subobjective of this cumulative dissertation is to obtain a set of relevant features for modeling the highly dynamic user-level behavior in heat networks. The parameters are to be selected to provide information on heat network efficiency and flexibility. As feature engineering and selection are time-consuming processes, this cumulative dissertation aims to accelerate model building by identifying suitable inputs for a data-driven model of parameters reflecting the highly dynamic user-level behavior, thereby enhancing the applicability of ML in heat networks.

The publication associated with subobjective 2 proposes the supply and return temperatures, as well as the volume flow at the user level as modeling targets that convey relevant information on heat network efficiency and flexibility. For these parameters, feature engineering and selection are conducted. A methodological investigation and selection process is developed. The transfer of influencing factors from heat-load prediction to user-level parameter modeling is evaluated. Pursuing subobjective 2 provides a set of relevant features for the suggested user-level modeling targets.

### 2.2.3 Subobjective 3: Dynamic User-Level Parameter Modeling in Heat Networks with Deep Learning Long Short-Term Memory

Subobjective 3 is to develop an ML-based modeling framework for the suggested parameters that reflect the highly dynamic user-level of the investigated heat network. The further aim is to reduce modeling error to an acceptable level by thoroughly optimizing the modeling setup. As existing knowledge of ML-based user-level modeling in heat networks is scarce, publication 3 integrates findings from subobjectives 1 and 2 and explores the knowledge required to establish an ML-based modeling framework for selected parameters that reflect highly dynamic user-level behavior.

To achieve this, an ML modeling architecture is first selected. Then, data pre-preparation is addressed, as it has been shown to significantly influence the success of modeling [63]. This includes, e.g., equalizing timestamps, cleaning outliers, and filling missing data [10]. Furthermore, this process involves encoding non-numerical information (e.g., human behavior [15, 16, 34, 40, 56, 65]), scaling [63, 76], and data transformation [20, 47], which are assessed. Moreover, an ML model does not perform optimally without optimizing the hyperparameters [43]. However, research on the optimal hyperparameter selection for modeling heat network parameters is limited [76]. Methods such as Bayesian Optimization allow the transfer of information on suitable hyperparameter ranges from comparable modeling tasks [37]. Therefore, suitable hyperparameter intervals to be tested through an optimization process are determined in this cumulative dissertation. Furthermore, domain-specific factors affecting modeling performance include seasonality [2, 7, 40] and the amount of data collected [7], which are therefore investigated as well. The named tasks are addressed to optimize a model for the highly dynamic user-level behavior in heat networks.

Particularly, publication 3 optimizes the modeling framework to reduce the error margin to an acceptable level despite the high dynamics of user-level behavior. The gained information contributes to accelerating ML applications in subsequent work. The findings motivate further investigation into integrating information derived from modeling and analysis of highly dynamic user-level behavior into heat network control to enhance the efficiency and flexibility of future heat networks.

### 3 Overview of Publications

The following sections present the results of the three respective subobjectives, citing the associated publications of this cumulative dissertation and the supplementary materials. A brief introduction to each publication provides context. Moreover, the author contribution statements are presented separately.

#### 3.1 Publication 1: Clustering- and Statistic-Based Approach for Detection and Impact Evaluation of Faults in End-User Substations of Thermal Energy Systems

Subobjective 1 is to enable an automated filtering of datasets that are unsuitable for modeling, e.g., because of the occurrence of faults in the substations that provide the recordings. In practical implementation, publication 1 aims to enhance automation in fault management for user-level substations in heat networks. Effective fault management is an essential step toward achieving the required transformation for heat defossilization. Yet existing fault detection processes exhibit shortcomings that strain operators' resources for fault management. Hence, increasing automation in this context is intended to advance fault management at the substation level, thereby enhancing heat network efficiency and flexibility.

Overall, the findings of publication 1 indicate that substations' primary-side temperature data are suitable for detecting user-level faults. Using a clustering approach, three indicators of existing faults were identified and labeled with expert knowledge: An exceeded return temperature, low cooling, and an inversion of the temperature difference. The third indicator is yet to be applied broadly. The indicators are then formulated to be accessible by statistical analysis. Using basic principles of thermodynamics, the negative impact of the detected fault on the system is quantified by combining temperature and volume-flow information. As the requirements for establishing the fault management approach are defined in publication 1, the transfer of this approach to other model regions can be investigated to improve the efficiency and flexibility of heat networks.

Moreover, within the research design of this cumulative dissertation, publication 1 presents an automated method for filtering outlier datasets that are unsuitable for data-driven modeling, which is essential for improving data quality in the modeling framework developed in publication 3.

##### 3.1.1 Author Contribution Statement of Publication 1

S.A.W. carried out the conceptualization, formal analysis, and investigation, and prepared the original draft. S.A.W. and U.R.-B. developed the methodology. M.F. contributed to the formal analysis and investigation. D.V. acquired the funding with the support of S.A.W., M.F. and D.V. were in charge of supervision. All authors reviewed the manuscript. All authors have read and agreed to the published version of the manuscript.

##### 3.1.2 Publication 1

WEBER, S.A./FISCHLSCHWEIGER, M./VOLTA, D./RIECK-BLANKENBURG, U.: Clustering- and statistic-based approach for detection and impact evaluation of faults in end-user substations of thermal energy systems. Sci. Rep., 14, 32166, <https://doi.org/10.1038/s41598-024-82103-5>, 2024.

# scientific reports



OPEN

## Clustering- and statistic-based approach for detection and impact evaluation of faults in end-user substations of thermal energy systems

Samanta A. Weber<sup>1,2✉</sup>, Michael Fischlschweiger<sup>2</sup>, Dirk Volta<sup>1</sup> & Ulf Rieck-Blankenburg<sup>3</sup>

In response to climate change mitigation efforts, improving the efficiency of heat networks is becoming increasingly important. An efficient operation of energy systems depends on faultless performance. Following the need for effective fault detection and elimination methods, this study suggests a three-step workflow for increasing automation in managing defective substations on the user level within heat networks. The work focuses on a model region in northern Germany. The local heat network provides data in roughly hourly intervals, including the supply and return temperatures and the volume flow of the substations. Firstly, this study identifies common indicators of faults using k-means clustering analysis of the temperature data and expert knowledge: an exceeded return temperature level, very low cooling, and inverted temperature readings. With these indicators, the subsequent statistical identification approach confirms the successful detection of affected substations, with common diagnoses including disabled return temperature limitation units, defective motoric valves, and faults in the storage control. Lastly, the study evaluates the impact of faults on the system efficiency. Combining the temperature and the volume flow data, the workflow quantifies the negative influence of a fault, enabling the prioritization of fault elimination measures in practical application to enhance the overall system efficiency.

**Keywords** Heat network efficiency, K-Means clustering, Fault management, Thermal energy systems

Heating systems are widely considered the most suitable solution for areas of high heat demand density<sup>1–3</sup>. However, transitioning existing systems to the 4th generation of district heating, which operates at lower temperatures, poses significant challenges<sup>4</sup>. Effective methodologies are required to reduce operating temperatures while accounting for existing network infrastructure.

Substations, which transfer heat from the network to users, are critical components regarding the heat network efficiency and often exhibit faults<sup>5–8</sup>. Faults in the substations typically require increasing either the supply temperature or the volume flow<sup>9</sup> through the substation, counteracting efforts to reduce system temperatures. Fault detection at the user level is crucial for enhancing thermal energy systems' efficiency, as single users can significantly impact the overall network performance<sup>10</sup>. Effective fault management could achieve half of the required temperature reduction for the transformation to 4th generation systems<sup>8</sup>. Besides improving efficiency, fault detection enhances reliability and safety<sup>11</sup> and can achieve cost savings<sup>12</sup>. Thus, developing feasible fault detection methods for user-level substations is essential.

Fault detection methods are mostly data-driven as they typically require large quantities of measurement data<sup>13</sup>. These include supervised (regression, classification) and unsupervised (e.g., clustering) models<sup>12</sup>. The standard strategy of fault detection methods is to compare the measured state of the system with an expected, faultless state<sup>13</sup>. With the rise of smart meters, these approaches have become increasingly relevant<sup>14</sup>, yet

<sup>1</sup>Energy and Life Science, University of Applied Sciences Flensburg, 24943 Flensburg, Germany. <sup>2</sup>Chair of Technical Thermodynamics and Energy Efficient Material Treatment, Institute of Energy Process Engineering and Fuel Technology, Clausthal University of Technology, 38678 Clausthal-Zellerfeld, Germany. <sup>3</sup>Business Segment Networks, Stadtwerke Flensburg GmbH, 24939 Flensburg, Germany. ✉email: samanta.weber@hs-flensburg.de; samanta.weber@tu-clausthal.de

existing methods face challenges like limited scalability, significant manual intervention, and complexity<sup>5,9,14,15</sup>, counteracting the practical application.

The literature emphasizes the need for effective fault detection<sup>5,8,9</sup>. For instance, Månsson et al.<sup>9</sup> underline the relevance of a well-operating model for fault detection, while Buffa et al.<sup>5</sup> highlight the advantages of machine learning despite its high data requirements. Manual fault detection is time-intensive<sup>5</sup>, and substations with low overall heat demand are often assigned lower priority or not monitored at all<sup>9</sup>. Yet, aiming for lower supply temperatures, this currently existing fault tolerance<sup>8</sup> should not be accepted but regarded as an optimization potential. In a related context, Fabre et al.<sup>17</sup> focus on improving user heating systems by reducing return temperatures. Given the potentially large number of substations, Ref<sup>17</sup> emphasizes the importance of prioritizing those with the most significant negative impact. For this, a limited number of operators consider the heat or volume flow of the station<sup>7</sup>. Thus, fault detection methods should be cost- and data-effective enough to consider all substations in a system rather than only those with the highest heat demand and quantify the negative impact of the fault on the heating system.

Further studies emphasize the need for reliable, scalable, cost-efficient, and simple solutions that minimize manual processing<sup>14,15</sup>. Data accessibility is one primary inhibitor, e.g., for secondary-side fault detection<sup>16</sup>. The objective must be to use easily accessible data (e.g., primary-side data, not secondary-side data, often challenging to collect) to simplify fault diagnosis and minimize manual intervention<sup>15</sup>.

Automation is essential for effective fault management but requires increased efforts in data labeling<sup>12,19</sup>. A majority of studies focus on anomaly detection with unsupervised learning because labeled data is scarce<sup>12</sup>. As faults are a subgroup of anomalies, existing approaches can lead to an overextension of the resources an operator can invest<sup>5,12</sup>. Ref<sup>12</sup> concludes that an effort should be put into further advancing actual fault detection alongside anomaly identification.

Clustering is a commonly applied method for grouping data and, thus, generating labels. Calikus et al.<sup>10</sup> and Gianniou et al.<sup>18</sup> introduce clustering methods to automatically identify user data patterns in heat networks to increase knowledge of thermal energy systems. Ref<sup>10</sup> emphasizes the relevance of continuous monitoring for detecting errors. The approach allows experts to outline substations that do not meet the behavior expected for the assigned control strategy<sup>10</sup> and, thus, aids in labeling data.

However, there is no standard monitoring approach to fault detection on the substation level<sup>7</sup>. Ref<sup>7</sup> identifies monitoring the return temperature, the temperature difference between the supply and return temperatures, and consumption or over-consumption (measured supply in comparison to an expected value) as existing approaches for fault detection. Another existing label for fault detection is the loss of user comfort<sup>21,22</sup>. Certain studies define a fault as always affecting user comfort and assign the term anomalies to the remaining phenomena<sup>21,22</sup>. Others, e.g., Ref<sup>7,17,20</sup>, express that inhabitants often do not notice faults in substations if no loss of comfort occurs. However, independent of the chosen terminology, the impact on user comfort can be insufficient as a criterion for fault detection. Therefore, methods for fault detection should be more standardized.

Certain studies showcase the application of the time-series data collected for the temperatures and the volume flow for fault detection. In the practical application of fault diagnosis, time dependencies, e.g., strongly relying on the information in the sequential order of time-series data, inflict an additional dimension and, hence, increase the complexity of the approach. In the context of reducing time dependency, Gadd et al.<sup>6</sup> suggest the “excess flow” (increased volume through substation caused by a fault) and the “thermal signature” (correlation of the temperature difference achieved by the substation and the outside temperature) as indicators of substation performance. In a subsequent study, Gadd et al.<sup>8</sup> inquire about the nature of occurring faults in user substations, found to be unsuitable heat load patterns, a low average annual temperature difference, and poor substation control. Calikus et al.<sup>20</sup> establish a method for fault detection working with the so-called “heat power signature” (or “energy signature”, the correlation of heat demand and outside temperature<sup>6,8,15,20,23</sup>) for the detection of errors in heat networks. The studies<sup>6,8,15,20,23</sup> suggest that relevant information for fault detection can be obtained with limited consideration of the time dimension. Thus, research should focus on investigating methods that reduce the dependence on the sequential order of time-series data and develop strategies to simplify data pre-processing without losing relevant temporal information.

An additional benefit of fault management is increasing measurement data quality for research purposes. Ref<sup>24–27</sup> highlight the importance of high-quality data for advancing heat network models. Improving fault detection directly increases network efficiency. However, it also enhances data quality and, thus, indirectly supports more advanced data-driven system modeling, which is considered highly relevant for further improving system efficiency<sup>28</sup>.

This study aims to develop an effective workflow for using substation data for fault detection on the user level in heating systems, with a focus on achieving a high degree of automation. The objectives of effectiveness and a high degree of automation shall be assessed by the following:

First, the workflow should reduce manual intervention. This study aims to contribute to this objective by utilizing available primary-side system data and encoding the employed data pre-preparation methods for automation. Also, the number of substations to be manually assessed shall be reduced, and the detection of faulty substations shall be fully automated.

Second, the method should lessen the reliance on the sequential order of time-series data, ensuring a more flexible and efficient process. Therefore, fault detection shall occur based on the sample distribution rather than by exploiting the information that the sequential order of the data can provide.

Third, all substations, regardless of their respective size, shall be monitored. To achieve this, the faults' impact on the heating network shall be quantified, allowing for the prioritization of elimination measures. The impact quantification shall also be fully automated.

The article is structured as follows: Section “Materials and methods” outlines the research methodology. The study introduces the framework of the fault identification methodology, presents the model region, and

provides the measurement architecture for data collection. Additionally, this work demonstrates the concept of the temperature signature of substations and introduces the necessary steps for data pre-preparation to use the k-means clustering approach. Further, the statistical identification process, as well as the quantification of the faults' impact, are explained. Section "Results" presents the results of the case study, including the identification and labeling of indicators with clustering and expert knowledge, fault detection with statistical methods, and the impact quantification of the faults. Furthermore, this work presents certain special cases of detected temperature signatures. Section "Discussion" provides the discussion. Section "Conclusions" finishes with the conclusions.

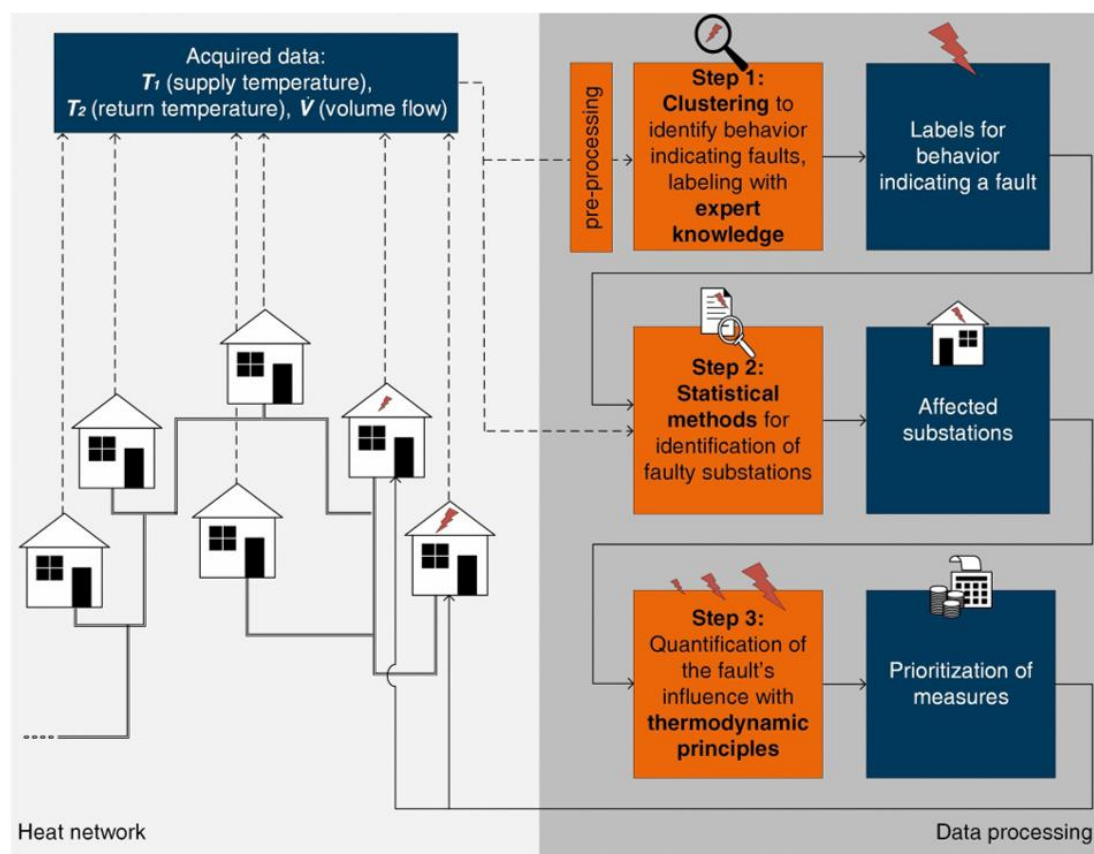
## Materials and methods

### Fault detection framework

In this work, fault detection occurs in a three-step workflow, highlighted in Fig. 1. First, typical patterns of substation data indicating faults are determined and labeled with clustering and expert knowledge. Second, a statistical approach for detecting faulty substations, automated by encoding, is developed. Third, the same code assigns a numerical value to quantify the fault's impact on the overall system.

Figure 1 illustrates the processing steps in orange and the obtained information in dark blue. The applied methods are highlighted in bold. As visible, the acquired raw data of the substations in the model region is pre-prepared prior to subsequent processing. As this study focuses on the interdependencies of the user substations and the heating system, we use the terms "fault" and "error" to describe substation behavior that has a disadvantageous impact on the heat network efficiency.

The pre-prepared data is then input into the workflow. For the first step, this study uses k-means clustering to establish the common (faultless) pattern as the reference for identifying anomalous observations in the temperature data of the substations in the model region. To achieve transferability, this work takes advantage of



**Fig. 1.** Fault identification workflow. The fault identification occurs in three steps, namely identification of the relevant faults with clustering, providing groups of faults to be labeled with expert knowledge, subsequently establishing statistical features to automatically detect substations affected by the labeled faults and assign a numerical value to quantify the (negative) impact of the faulty substation on the system to prioritize measures with thermodynamic fundamental relationships.

the statistical nature of k-means rather than explicitly defining a specific faultless reference pattern. Clustering-based methods are unsupervised and can identify patterns among data and assign subsets with similar patterns to groups. This step might require comparatively high computational resources. As the clustering provides no reason for the decision, expert knowledge is required to provide labels for the groups. It is used to label statistically accessible indicators to detect defective substations. In this study, patterns indicating faults are labeled with statistically accessible features.

In the second step, these indicators can be used to identify faulty substations automatically through an encoded statistical process. The acquired labels and the raw data are the inputs. Such a statistical evaluation typically requires less computational effort than clustering. However, this study's method demands a feature indicative of an error, which the clustering approach provides. The statistical approach returns the substations affected by at least one of the faults labeled in the first step.

Thirdly, information on the intensity of the negative impact of the detected faulty behavior on the heating system can be provided, relying on the thermodynamic relationships of volume flow and temperatures. The information on the occurrence of a fault obtained in the second step is insufficient, as it does not provide any measure of prioritizing elimination actions. This can overstrain the operator's resources for fault management. Hence, the code automatically returns a numeric quantification of the fault's impact on the system.

Automation enables repetitive conduction of fault detection and impact evaluation. Therefore, conducting the second and third steps of the suggested workflow at regular intervals using the recently recorded data outlines existing, newly occurring, and reoccurring faults.

### Model region and data acquisition

This study employs exemplary data of the supply and return temperature as well as the volume flow recorded in roughly hourly time steps between 1<sup>st</sup> January and 31<sup>st</sup> December 2023, of a centrally supplied district heating system in Tarp, northern Germany, with 486 substations. Figure 2 a) displays the general structure of the system, Fig. 2 b) shows the recorded values of the volume flow, the supply, and the return temperature for a day of January (mid-heating season) and July (mid-summer) for one exemplary building in the heating system. Figure 2 c) depicts the infeed heat flow for the same days and Fig. 2 d) the infeed heat flow over the hour of the year.

For the infeed of the heat network, the supply temperature ranges between 70 and 105 °C. The substations are designed to receive the minimum supply temperature of 70 °C for an outside temperature  $\geq 15$  °C and up to 100 °C for an outside temperature  $\leq -10$  °C. Currently, the return temperature received at the infeed plant reaches values between 50 and 55 °C, which is a comparatively high value compared to more modern heating systems.

The heat network mainly covers the heat demand of individual and multi-family homes. Certain buildings share a common substation. One substation supplies a living quarter (see the upper left corner of Fig. 2 a). One industrial customer and buildings used by the public (e.g., schools, gyms, churches) are connected to the system as well.

The fault management method of this study rests predominantly on the data of the individual substations, as captured exemplarily in Fig. 2 b). Users in the model region can receive heat for domestic hot water and space heating or the latter only. Further, it is possible to heat the domestic hot water directly through the heat exchanger or use a storage facility, which impacts the temperature level required for the supply and the resulting return temperature. Figure 3 depicts a typical measurement architecture on the end-user level within the model region.

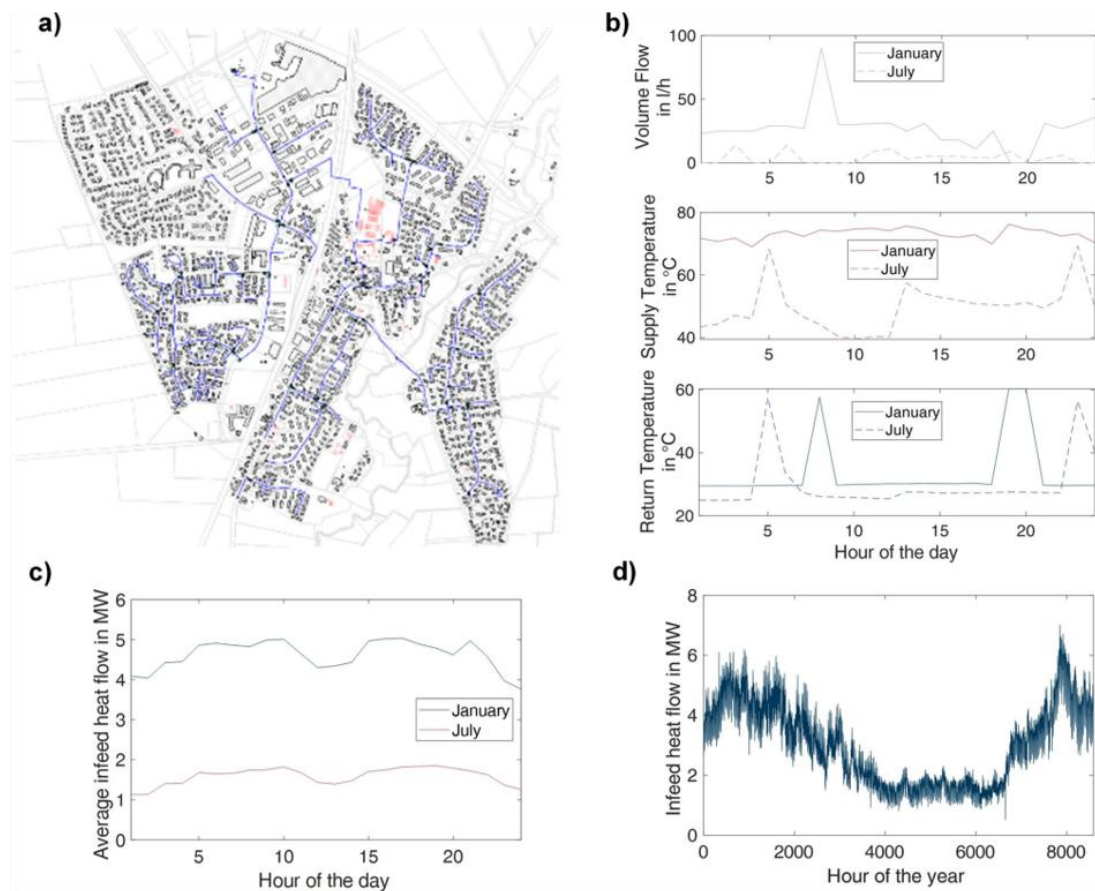
The measurement architecture captures the sum of the heat supplied for domestic hot water and space heating. The motoric valves, as actors, determine the volume flow values. As the scheme in Fig. 3 is an advantageous setup, it should be mentioned that the real-case scenarios can (strongly) differ, possibly influencing the efficiency. For the direct domestic hot water supply without a storage facility, a low flow rate of primary side heat carrier is required for times of no demand. This maintains the level of temperature needed when demand occurs again. Hence, in this specific case, no complete stagnation of the substation can occur, but the volume flow is limited to avoid excessively high return temperatures.

If measures for eliminating the faults lie in the user's sphere, the operator can only give advice. Ref<sup>7</sup>, therefore, suggests that a good relationship with the user is highly relevant in fault elimination. However, legally binding contracts containing threshold values to be adhered to can exist, as in the model region. Specifically, the return temperature maximum to be adhered to on the average is 55 °C. An excessive return temperature causes relevant heat losses in the return line. Also, it is indicative of a low cooling of the heat carrier, which occurs for values of the volume flow surpassing the optimum. Thus, the hydraulic load of the system is affected. Additionally, the losses in the supply line are increased: for a higher capacity flow (product of volume flow and heat capacity), the mean overtemperature of the supply line rises. Therefore, a user's substation with a high return temperature has a substantial negative impact on the system. The operator determined that the boundary of 55 °C is the maximum level to limit the network losses to an acceptable level. As the user substations are the property of the user and also the user's responsibility, the restriction to 55 °C further aims to provide a legal foundation for the operator to oblige the user to take measures.

It should be acknowledged that the limit value of 55 °C is specific to the model region and can appear high compared to other heating systems. The value can be adjusted, e.g., for more modern and modernized systems, as well as systems conceptualized under a different legal framework.

### Temperature signature of substations

In this study, following the terminology of Ref<sup>20</sup>, the term "temperature signature" refers to the supply temperature measurements of a substation displayed over the return temperature readings. Similarly to, e.g., Ref<sup>6,8,20,23</sup>, this study does not primarily rely on the information in the sequential order when treating the time-series data.



**Fig. 2.** Network structure of the model region, data of one substation and for the infeed heat flow for a day in January and July, and dynamic data of the infeed heat flow over one year. The model region in **a)** is a centrally supplied district heating network with predominantly stub lines and one closed circle line; the data in **b)** visualizes the dynamics of one substation over a day; **c)** highlights that the infeed heat flow over a day follows two demand peaks and the general level depends on the season, further visualized by **d)** the heat infeed over the year.

Instead, the data are investigated with regard to the sample distribution. As temperatures alone indicate the efficiency of a thermal energy system<sup>9</sup>, this work hypothesizes that information on the occurrence of faults can be captured without relation to time, employing the supply and return temperature data of user substations for the identification of faults. Working with limited time dependency serves the objective of avoiding time-intensive pre-processing and reduces computational costs.

To further explain the concept, Fig. 4 contains temperature signatures for two substations. On the left, the plots distinguish between the summer (1<sup>st</sup> June to 15<sup>th</sup> September 2023) and the heating period, and on the right, between the existing operational modes. This study considers return temperatures above 55 °C as disadvantageous for the system efficiency. On the contrary, the lower the return temperature drops below 55 °C, the stronger the advantageous influence becomes. A black solid line in Fig. 4 indicates the equality of supply and return temperature.

The distance in pipe length from the infeed strongly influences the temperature signature of substations. Aiming for an efficient heat supply, the most advantageous user is situated close to the heat and power plant<sup>29</sup> because the pipe length is minimal, and consequently, the heat losses are low. For this user, the supply temperature will range amongst the highest in the system. Such a situation is depicted in Fig. 4 a), b) (same substation as in Fig. 2 b)). On the other hand, if the user is remotely located (Fig. 4 c), d)), the heat carrier cools down in the distribution pipes and the connection pipe, which affects the temperature signature but does not necessarily indicate a fault.

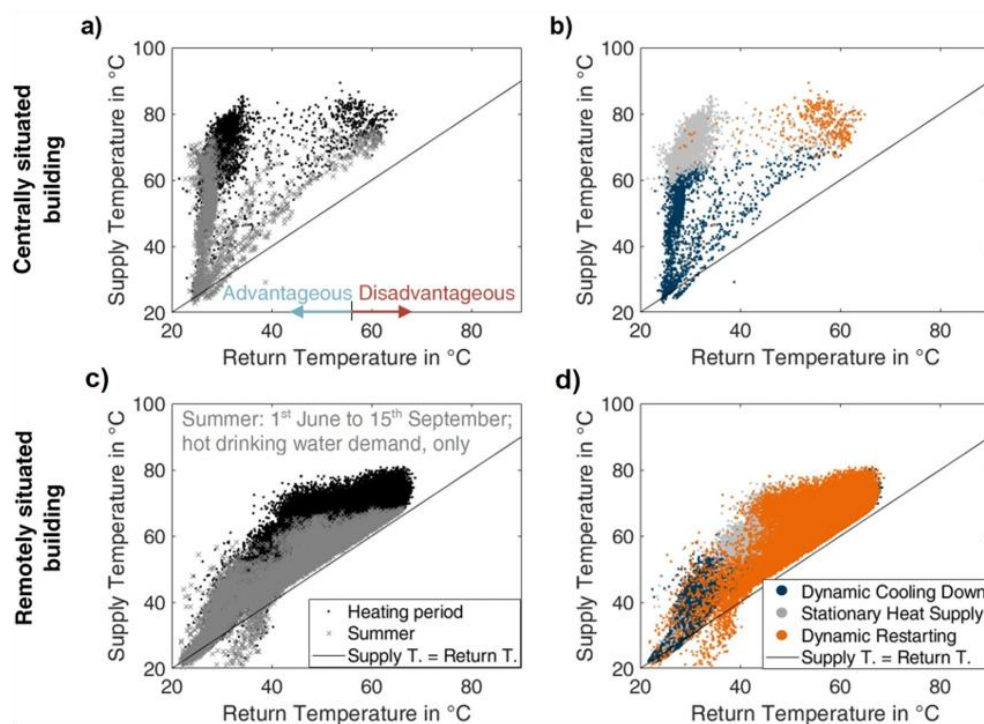
Figure 4 a) and c) distinguish the summer (periods of usually only domestic hot water demand) and the heating season (periods of both space heating and domestic hot water demand) to visualize the seasonal impact.

To account for seasonality, this study assumes that the data should be collected over at least one year, as a fault detection method should cover both periods. As the heat exchange for domestic hot water and space heating employs separate heat exchangers, faults can be categorized as affecting both heat exchangers or only one of the two by the distinction of the periods. Additionally, the point of time the fault occurred can be narrowed down. This kind of visualization shall, therefore, be applied in the following.

- If the substation's architecture allows for stagnation (zero volume flow, not possible for direct domestic hot water heating), the facility cools down (Fig. 4 b), d), blue dots) for periods of no demand, which is advantageous for the system. Stagnation occurs more often in summer (Fig. 4, grey crosses, a), c)), when domestic hot water is required only.
- In comparison, for periods of relatively constant heat demand, a state of approximately stationary heat supply will be reached (Fig. 4, grey dots, b), d)) with intermediate values of the volume flow. The substation's control should adjust the volume flow to achieve the highest possible cooling of the heat carrier. This reduces the required volume flow and the resulting return temperature.
- When restarting the facility, the volume flow strongly increases, resulting in disadvantageously low temperature differences (Fig. 4, orange dots, b), d)). The low temperature difference places the values associated with high volume flow values on the right in the plots in Fig. 4. Hence, for an efficient substation, a limited number of samples should fall into this mode.

Measurements below the black solid line form a particular case caused by storage capacities in the system, and there are two possible explanations. Both rely on the fact that the return temperature measuring probe is often located physically closer to the heat exchanger:

- For a stagnating substation cooling down as in Fig. 4 a), b) (blue dots), the heat stored in the heat exchanger warms the return temperature measuring probe, while the supply temperature measuring probe (located further away) is unaffected.
- For a restarting facility like in Fig. 4 c), d) (orange dots), the cooled-down heat carrier in the distribution pipe must be replaced. A long stub lane connecting the substation to the system leads to a higher amount of water to be replaced. Until the hot heat carrier is received again, the heat carrier on the primary side can be warmed up by the secondary side heat carrier from the heat storage facility.



**Fig. 4.** Temperature signatures of end-user substations outlining the line of equality for supply and return temperature. The substations in **a**, **b** are centrally and in **c**, **d** remotely positioned, and the diagrams enhance **a**, **c** summer (domestic hot water demand, only; 1st June to 15th Sept., grey crosses) and heating period (space heating and domestic hot water demand, black dots), and **b**, **d** the general operational modes existing for faultless behavior distinguished by the value of the volume flow (blue: zero for stagnation, grey: intermediate for stationary supply, orange: high for restarting), advantageous return temperatures defined as below 55 °C; the black solid line indicates equality of temperatures.

As can be seen, the operational mode influences the position of data points in the temperature signature. The time dependencies are related to the number of samples. Both combined lead to an accumulation of data points in the temperature signature, which is, therefore, relevant information that shall be used for clustering. To establish labels for faulty behavior, the faultless patterns must be known, which can range between the two extremes of locations depicted in Fig. 4.

#### Clustering and expert knowledge for identification and labeling of indicators for faulty behavior

Clustering is an unsupervised machine learning technique<sup>10</sup> that aims to identify subgroups of data that differ from each other as strongly as possible but contain subsets of data that are as similar as possible<sup>30</sup>. This study employs the principle of k-means clustering. It splits a dataset of  $i$  observations into  $k$  clusters using two conditions<sup>30</sup>:

- The centroid of every cluster is the mean value of the observations in the cluster.
- Every observation is assigned to the cluster with the nearest centroid.

The algorithm itself contains five steps<sup>30</sup>:

1.  $k$  observations are chosen randomly as initial centroids from the dataset.
2. The distance of the remaining observations to the initial centroids is calculated.
3. Every remaining observation is assigned to the cluster with the nearest centroid.
4. The mean value of the newly established cluster is treated as new centroid.
5. Steps 2 to 4 are repeated until the algorithm converges.

The substations' datasets form separate observations. As given, one observation is automatically chosen as the initial centroid for each cluster. Thus, applying the method more than once leads to differing results. This study

takes advantage of this characteristic. Further, an ML method is unbiased and objective. This ensures that no subjective considerations of a (human) expert influence the results.

Data pre-preparation, i.e., feature engineering, must occur to enable comparability of the different substations' datasets. Not all observations show equal numbers of measurements, so the number of data points in the observations must be equalized. Further, as described, the data accumulation in the temperature signatures fuses the information on the temperature, the volume flow, and the time dependencies and, hence, forms a relevant feature for evaluating a substation's efficiency.

Therefore, to achieve comparability and integrate all the information, this study employs the representation of the temperature signatures as heat maps, which visualize data accumulation. For the processing steps in Fig. S3 (Supplementary Information), the scatter plots of the substations, as in Fig. 4, are regarded as grids with steps of 1 K in the two temperature directions. Evaluating the share of the overall number of samples falling into a specific grid position provides information for clustering.

Fig. S3 a) (Supplementary Information) shows the steps to obtain a heat map that represents data accumulation from the temperature signature. The modeling environment in this study is Matlab. As the predefined k-means clustering algorithm<sup>31</sup> demands one-dimensional datasets, the two-dimensional datasets are parallelly transformed into single-row vectors by concatenating the rows of the grid from the lowest to the highest supply temperature, as visualized and described by the flow chart in Fig. S3 b) (Supplementary Information).

This step enables the comparison of observations with differing data collection rates as long as the recordings capture the relevant information on data accumulation. In contrast, limited numbers of collected samples, in particular, result in accumulations at specific points. These observations should be excluded. As a comparatively large number of substations achieves 90% of data collection, this study employs this share of 167 sets of end-user data for the analysis. For data privacy reasons, no traceable indicator identifies the individual substations.

The subsequent step to data pre-preparation is setting the clustering approach's hyperparameters, in this case, the target number of clusters. As shown in Fig. 4, two general, expectable (faultless) behavioral patterns of substations exist: the advantageous and the rather disadvantageous patterns with regard to their influence on the system. The expectation is that these will form two clusters with a large number of assignments based on the expected (faultless) behavior.

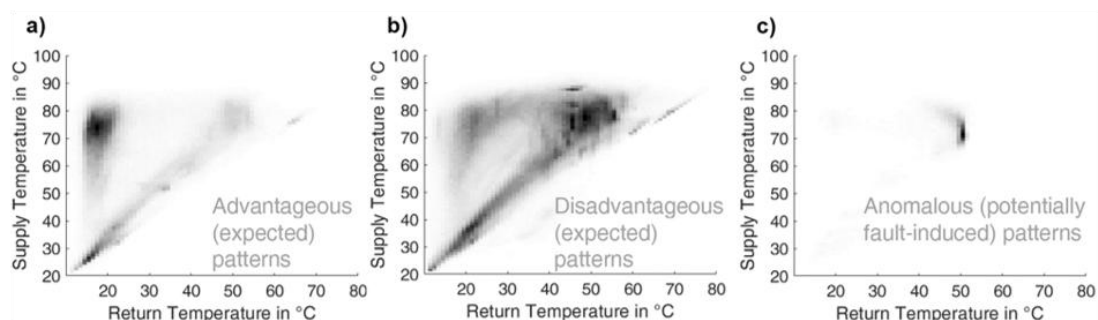
Therefore, the hypothesis is that if the target number of clusters is set to three, this third cluster should contain only a small subset of the overall number of observations, most likely patterns similar to each other but differing strongly from those of the first two clusters. This study hypothesizes that the faulty and anomalous substation patterns will be found in this third, smaller cluster.

To identify all of these patterns present in the dataset, all substations prone to being assigned to the third cluster must be determined. This study takes advantage of the fact that the initial centroids of the clusters are chosen statistically by k-means. When the clustering is conducted repetitively for a sufficient number of times, all anomalous patterns will eventually form the third cluster at some point in the repetition.

Furthermore, this study relies on the circumstance that differing patterns should appear only for a small number of observations. Therefore, if the third cluster contains a high number of substations, the anomalous patterns are not successfully identified. With iterative manual testing, a suitable maximum number of assignments for a "small" cluster for the specific case of the model region's data is determined to be five.

Apparently, certain anomalous patterns occur for a group of substations. If one of these forms the centroid of the third cluster, all similar patterns are assigned to this group. Gradually increasing the number defined as "small" from one to four leads to the detection of a small number of substations with indeed anomalous but unique patterns. For the number of five, considerably more substations with anomalous patterns are detected. Increasing the number of clusters further, a growing number of substations with only slightly anomalous patterns form clusters.

Figure 5 displays the centroids of the resulting clusters for an exemplary execution of the process of splitting the data into three groups. To display the centroids, the single-row vectors returned by the k-means algorithm are reordered to the initial matrix shape of the 1-K-grid. Displaying the matrix as a heat map, lighter colors in



**Fig. 5.** Heat maps of exemplary centroids of the clustering process with 3 clusters. Clusters **a)** advantageously operating substations and **b)** more disadvantageously operating substations contain the general behavior, while **c)** deviating operational patterns indicate anomalies.

Fig. 5 indicate lower ratios of the collected data falling into the grid position. Hence, the darker the color in a specific grid position, the more data were recorded in this grid position for the assigned substations.

Visibly, the first cluster contains the more advantageously operating substations (tendency to left in heat map of temperature signature), the second the more disadvantageous (tendency to right in heat map of temperature signature), and the third a share of substations with anomalous patterns, as expected. Similar results are returned for repetitive conduction, where the substations assigned to the third cluster change. This study, therefore, concludes that splitting the observations into three clusters serves the objective of outlining the differing temperature signatures but requires repetitive execution to detect all existing anomalies.

To ensure all anomalous patterns are identified, this work executes the clustering process 10,000 times. For each substation, we calculate the number of times it is found in a cluster with a maximum of five assignments. The choice of 10,000 executions relies on the circumstance that the probability of the substations being in a small cluster becomes relatively stable when statistical effects level out. Further increasing the number of executions has a neglectable impact on the results. The number of times a substation is assigned to a small cluster indicates the degree to which the pattern differs from all others ("degree of anomaly"<sup>20</sup>). Figure 6 shows a bar graph for each substation with at least one assignment to a small cluster in descending order. The substations with the patterns with the highest degree of anomaly are found on the left.

The clustering assigns 95 of the 167 substations to a small cluster at least once. Therefore, clustering enables the objective reduction of the number of patterns that need to be manually investigated. The patterns that differ the strongest from the remaining dataset are assigned to a small cluster in up to 15% of the executions. It should be considered, as underlined in Ref<sup>47</sup>, that anomalous behavior is not directly an indication of a fault and that, if indeed a fault exists, the specific cause must be determined. Hence, for this study, all 95 of the substations' patterns identified as anomalous were manually investigated with expert knowledge to obtain labels for the most relevant existing patterns indicative of the occurrence of a fault.

### Statistical identification of faulty substations and influence quantification

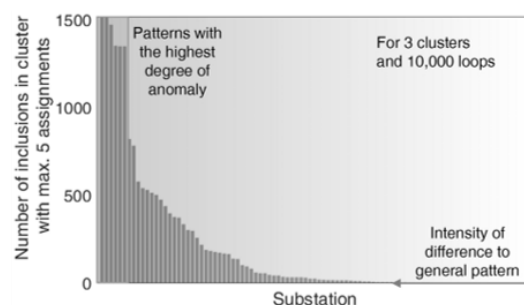
Statistical analysis is a method conducted traditionally<sup>7</sup>, as it is a computationally efficient option for fault detection. Additionally, as a physical representation of the indicator exists, the method is comprehensive. However, statistically accessible labels are required. Hence, it can be conducted only once labels for fault indication are established.

In this study, the application of the statistical fault identification thus depends on the labels acquired with the clustering approach and can only be conducted subsequently. Unlike the clustering approach, the statistical evaluation shows low sensitivity to the rate of collected data. Thus, with this second step of the statistical fault detection process, all 486 substations in the heating system can be covered.

Further, for an operator with limited resources, more than detecting the faults is required. The greater a fault's influence on the system, the more critical eliminating the cause becomes. Historically, this has led to monitoring the substations supplying the most heat, possibly leaving faults in substations with low demand undetected<sup>9</sup>. Thus, all substations should be monitored to improve the system's general operation. Evaluating the negative impact of the faulty substations on the system increases the benefit because the operator can prioritize fault elimination measures starting with the most disadvantageous substations.

Thus, the fault's impact on the system should be quantified. This study employs the volume flow data recorded parallelly to the temperature data to reach this objective. The higher the volume flow, the stronger the impact on the system becomes.

Fig. S4 (Supplementary Information) provides a flow chart of the processing steps for the statistical fault detection and impact quantification. First, the indicators for the occurrence of a fault established in the first step of this three-step method are tested for each substation dataset individually. Second, it can be required to assess whether additional boundary conditions occur, which can be individual for the indicators (e.g., enough data collected, the general temperature level, etc.). Third, the heat transferred for the occurrence of the indicator is calculated as a quantification of the faults' impact. Additionally, it is suggested that a substation is only marked as faulty if the calculated heat exceeds a threshold specific to the indicator and the model region. This maintains



**Fig. 6.** Number of inclusions of a substation into a small cluster. The number is indicative of the degree of deviation of the related temperature signature from the general pattern, with small cluster meaning max. 5 assignments; results for 10,000 executions of k-means clustering with a target value of 3 clusters.

relevance and further narrows the meaningful information to the faults with the highest impact. The code automatically returns the relevant information and sorts the substations detected as faulty by the calculated heat values in descending order. Once the indicators are provided by step 1 of this study's three-step method, the general framework can be explicitly encoded for steps 2 and 3.

## Results

### Clustering and expert knowledge-based identification of indicators for common faults in use substations

In this first step, this study aims to identify the most common behavioral patterns indicative of faults in the context of labeling data with the clustering algorithm.

Figure 7 shows the temperature signatures of six of the ten substations, which are identified most often as providing unusual patterns (Fig. 6). The choice contains the most informative plots to cover all of the behavioral patterns indicative of faults identified as occurring most often.

The displayed datasets in Fig. 7 tend to exhibit accumulation in the upper right corner, which is disadvantageous for the reason of high return temperatures. As a legally binding return temperature limitation should strictly avoid more than 55 °C, each substation appears to show a fault leading to this behavior. Figure 7 a) and b) are neighboring substations of buildings. As Fig. 7 a) and b) appear to have similar temperature signatures, similar faults could have occurred. In the following, expert evaluation shall provide possible causes.

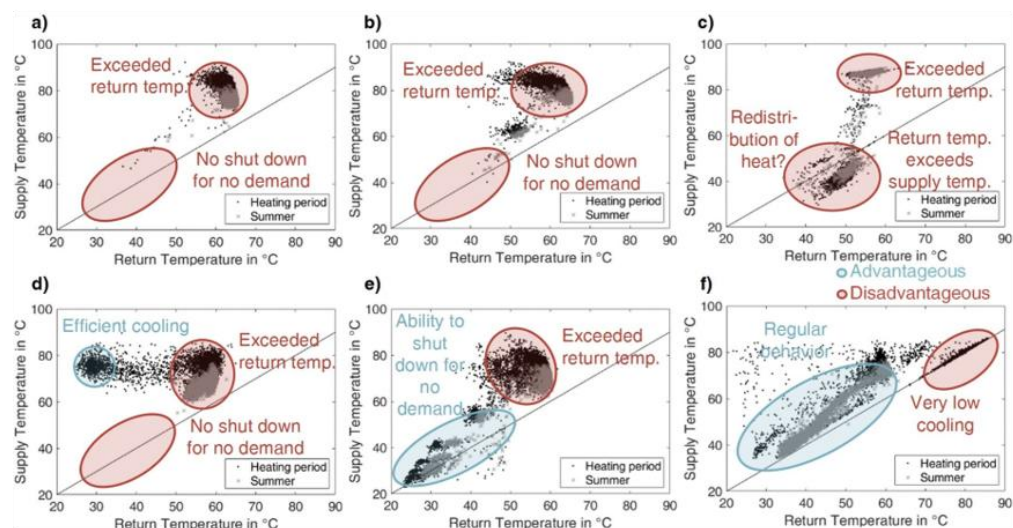
Figure 7 a) and b) exhibit very low cooling of the heat carrier. The control strategy or settings likely do not allow for a more advantageous operation.

For Fig. 7 c), many samples appear below the solid black line, indicative of recordings with a higher return than supply temperature. This behavior should occur only for a limited number of samples, which is equivalent to a low number of recorded hours for this state. Visits to the site revealed that a temperature signature like Fig. 7 c) indicates a faulty or at least disadvantageous control of the building's heat storage, possibly redistributing heat into the system.

The substation in Fig. 7 d) evidently could reach sufficient cooling, but this state occurs comparatively rarely and only in the heating season. If the supply temperature level were low, the inefficient operation could be related to the location in the network, as discussed above. However, regarding Fig. 7 a), b), and d), the supply temperature is sufficient, especially in the heating period, and never falls below 60 °C, even in summer. Therefore, it must be concluded that the substations Fig. 7 a), b), and d) show a constant flow of the heat carrier, possibly because of direct heat supply for domestic hot water, but evidently also associated with low cooling when demand exists. Hence, the control strategy or the facilities' architectures do not allow for higher cooling.

In contrast, Fig. 7 e) exhibits stagnation but appears unable to reach advantageous operating points when supplying heat, possibly for the same reasons as for Fig. 7 a), b), and d).

Figure 7 f) displays a particular case of behavior, hereby referred to as very low cooling. The volume flow is uncontrolled, and no heat is supplied for samples, indicating the same value for the supply and return temperatures (accumulation of data samples on the black line in the upper right corner). However, the fault



**Fig. 7.** Selection of 6 temperature signatures of substations with strong deviation from general pattern. (a) to (f) distinguish between the heating period (black dots) and summer (grey crosses) with an indication of advantageous (blue) and disadvantageous (red) positions of data samples and a reference to the impact on the heating system, where the 6 selected substations originate from the 10 substations identified as most strongly deviating.

leading to the very low cooling developed over time and during the heating period, which can be seen from the remaining data points. Yet, very low cooling should be distinguished from the behavior of a restarting facility, which can cause similarly low temperature differences but for lower temperature levels.

With expert knowledge, this study determined three prevailing types in the model region with the prevailing underlying causes in brackets:

1. Exceeded return temperature (disabled, modified, or non-existent return temperature limitation valve)
2. Very low cooling (defective or stuck flow control valve)
3. Inverted temperatures (defective or inadequate control of storage tank)

The very low cooling is a particular case of exceeded return temperatures, where the volume flow is not adjusted sufficiently but passes at a high rate in an uncontrolled or very limitedly controlled manner. Yet, as measures are often straightforward, a separate treatment of this behavior can be advantageous: As stated in Ref.<sup>7,17</sup>, the motoric valves are prone to faults. Often, blocking causes a very low cooling. Ref.<sup>7</sup> names the first two indicators as well. However, the third case could be identified based on the clustering approach as an additional indicator.

### Statistical identification of faulty substations and quantification of their impact on the heat network

Based on the labeled patterns indicative of faults acquired by the clustering approach, statistically accessible features for the identification of faulty substations are generated:

1. The mean return temperature reflects whether this value surpasses the threshold of 55 °C specific to the heat network under investigation. The higher the volume flow is for these readings, the higher the impact on the system.
2. Similarly, very low cooling can be identified by evaluating the temperature difference between supply and return temperature readings. As a boundary condition, the very low cooling is generally characterized by a high temperature level. Again, the volume flow readings enable quantifying the impact on the system.
3. Thirdly, the inversion of the temperature potential can be evaluated. For this case, a return temperature measurement higher than the supply temperature is most disadvantageous for the heating system if it occurs for readings of the volume flow > 0, as heat is redistributed in this case.

For an exceeded return temperature and very low cooling, this study quantifies the negative impact as follows. Firstly, only for a return temperature exceeding 55 °C on average does the operator have legal grounds to require measures. Employing this value as a reference is suggested. The mean value of the return temperature  $\bar{T}_2$  is formed and compared to 55 °C. The higher this temperature difference, the stronger the negative impact of the fault on the system. Therefore, the theoretical value for the surplus of heat  $Q_{55}$ , which should have been transferred in the substation to reach 55 °C in the return flow, is calculated (Eq. 1).

$$Q_{55} = m \times c_p \times (\bar{T}_2 - 55^\circ\text{C}) \quad (1)$$

$$m = \int_{t_{begin}}^{t_{end}} \dot{m}(t) dt$$

Calculating the inverse heat flow requires an additional step. Forming a mean value of the temperatures could lead to a leveling out of periods in which the potential has the expected direction ( $T_1 > T_2$ ) and periods of inversion ( $T_2 > T_1$ ). If the temperature readings indicate an inversion, but if no primary volume flow occurs, no heat is transferred. Thus, the returned heat is calculated according to Eq. 2.

$$Q_{ret} = \int_{t_{begin}}^{t_{end}} \dot{Q}_{ret}(t) dt \quad (2)$$

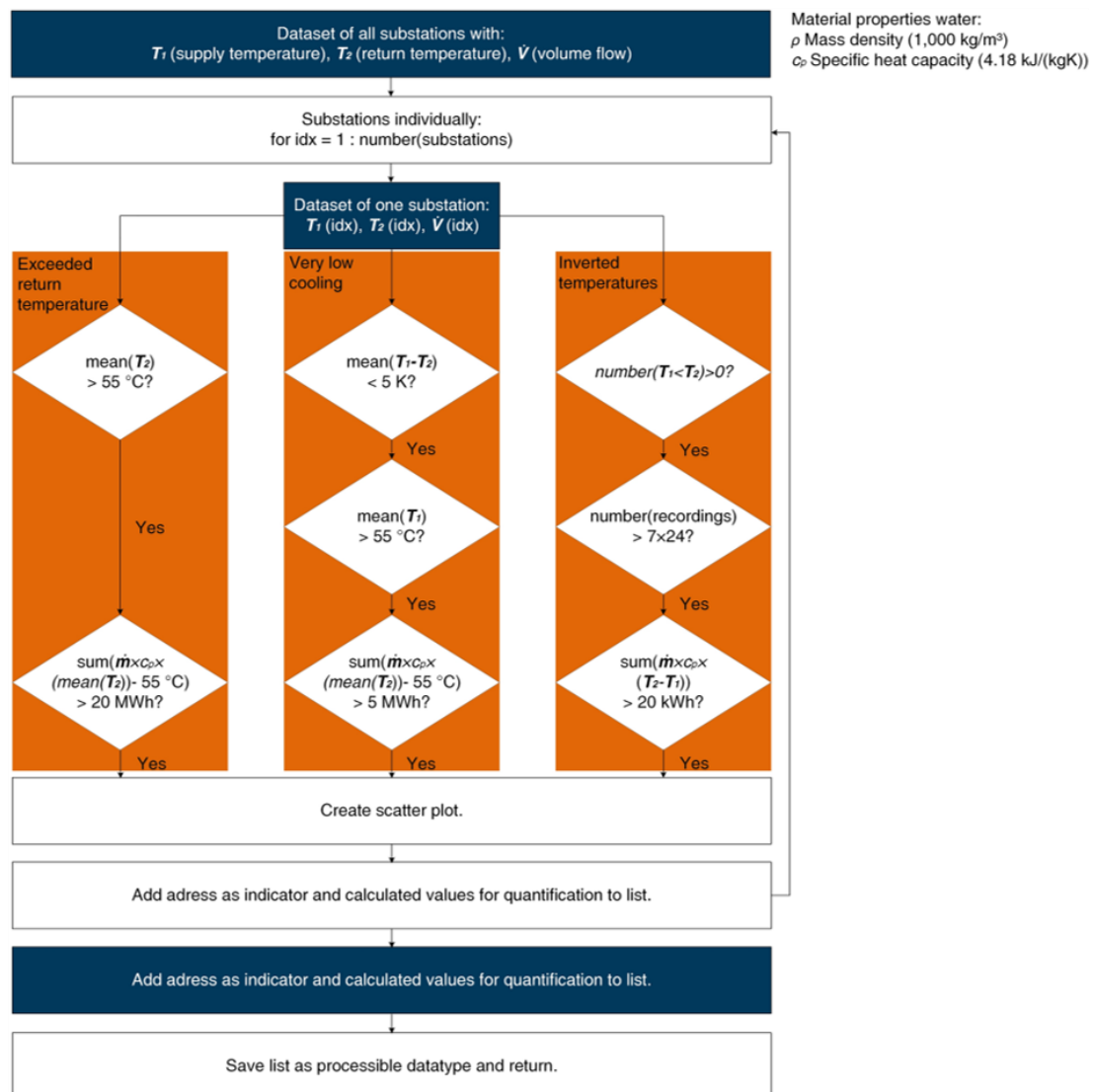
with

$$\dot{Q}_{ret}(t) = \begin{cases} \dot{m}(t) \times c_p \times (T_2(t) - T_1(t)) & \text{for } T_2 > T_1 \\ 0 & \text{for } T_2 \leq T_1 \end{cases}$$

Figure 8 outlines the explicit algorithmic flow chart for identifying faults and quantifying their impact resulting from the general framework in Fig. S4 (Supplementary Information). In the specific case of the model region, the integration is discretized in hourly timesteps, resulting in a summation replacing the integration.

Table 1 gives the results for executing the algorithm encoding steps 2 and 3 of the three-step method. The September 2023 data is employed exemplarily with the named thresholds and boundary conditions.

The assigned thresholds aim to maintain relevance regarding the operator's limited fault management resources. Conducting fault detection without any threshold reveals a leveling out of the numerical value assigned to the respective impact of the faults. In the case of the explicit sample set, 20 MWh of heat exceeding 55 °C in the return flow for the first fault, 5 MWh for the very low cooling, and at least 20 kWh for the returned heat could be identified as suitable with manual analysis. Below the thresholds, a saturation within the values in Table 1 starts to occur.



**Fig. 8.** Algorithmic workflow of statistical approach for the identification and impact evaluation of faults on the user level. The data are processed individually for each substation in three paths to evaluate whether a fault exists and determine the degree of influence on the system with the named restrictions, after which the results are visualized and saved in a user-friendly manner.

The returned heat threshold is specifically lower in order of magnitude than for the other two faults. However, detecting malfunctions in the substation control can have a strong influence: As a fault in the substation control usually leads to inefficient behavior, the return temperature can exceed the limits for times of high demand. In contrast, heat is returned in times of low demand. This means that, on average, the threshold of 55 °C might not be exceeded, but the impact can be high nevertheless.

It should be mentioned that evaluating the heat flow for all three cases is expedient only if sufficient samples are collected. Otherwise, the results returned for the discrete summation are, evidently, lower. The summation of returned heat has shown to be especially sensitive to reading errors for low amounts of collected data. Thus, imposing a boundary condition in this case and exclusively considering substation datasets with, e.g., at least one week of readings is suggested.

The substations in the “very low cooling” column of Table 1 relate to the numbers for exceeded return temperature when sorted in descending order. Thus, the numbers indicate that the approach for the return

| Results: exceeded return temperature                  |                               |                                    | Results: very low cooling |                                   |                                    |
|-------------------------------------------------------|-------------------------------|------------------------------------|---------------------------|-----------------------------------|------------------------------------|
| Substation                                            | Mean return temperature in °C | Surplus heat in return flow in MWh | Substation                | Mean temperature difference in °C | Surplus heat in return flow in MWh |
| 1                                                     | 61.41                         | 94.10                              | 2                         | 3.72                              | 30.05                              |
| 2                                                     | 68.16                         | 30.05                              | 5                         | 3.74                              | 8.06                               |
| 3                                                     | 66.10                         | 29.48                              | 6                         | 2.31                              | 7.86                               |
| 4                                                     | 56.12                         | 20.50                              | 7                         | 1.01                              | 5.36                               |
| Results: inverted temperatures and returned heat flow |                               |                                    |                           |                                   |                                    |
| Substation                                            | Returned heat in kWh          | Substation                         | Returned heat in kWh      |                                   |                                    |
| A                                                     | 151.50                        | J                                  | 29.29                     |                                   |                                    |
| B                                                     | 99.99                         | K                                  | 28.42                     |                                   |                                    |
| C                                                     | 67.64                         | L                                  | 25.66                     |                                   |                                    |
| D                                                     | 48.06                         | M                                  | 25.62                     |                                   |                                    |
| E                                                     | 44.76                         | N                                  | 25.24                     |                                   |                                    |
| F                                                     | 42.13                         | O                                  | 23.73                     |                                   |                                    |
| G                                                     | 41.42                         | P                                  | 22.71                     |                                   |                                    |
| H                                                     | 36.33                         | Q                                  | 21.79                     |                                   |                                    |
| I                                                     | 30.97                         | R                                  | 20.81                     |                                   |                                    |

**Table 1.** Results of statistical fault detection and impact quantification exemplarily for data of September 2023.

temperature does indeed include very low cooling, but it does not necessarily appear amongst the highest priorities.

The substations that returned noticeably high amounts of heat did not appear in any of the other two lists of faulty substations in Table 1, which is why they are referred to alphabetically.

For the investigated period of September 2023, counting without the application of the threshold, a considerable share of 49 (10%) of the investigated 486 substations exhibit an exceeded return temperature. Yet, only a limited number of these seriously impact the system. Still, the necessity for measures is emphasized by the results.

#### Special cases for temperature signatures related to faults and identified causes

Specific temperature signatures shall be addressed to outline the limits and chances of this study's methods. The objective is to enhance expert knowledge and clarify information that can be derived from a substation's temperature signature.

This study assigns causes to the visible temperature signatures in the context of labeling data. Subsequent research can aid in classifying faults automatically and reduce the necessity for manual processing even more. Also, the impact of the frequency of conducting fault detection becomes more apparent as the temperature signatures give indications of when the fault occurred.

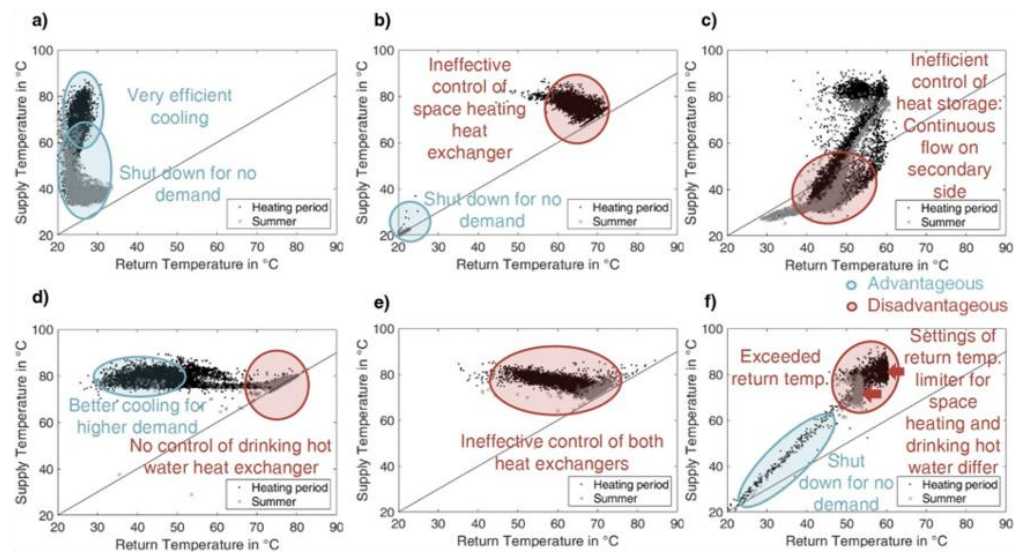
Figure 9 a) to f) show temperature signatures of substations identified as showing a differing pattern or as indicative of a fault.

Figure 9 a) gives an example of the most advantageous operation in terms of system efficiency. Visibly, the end-user's facility cools down very effectively, and the heat exchange stops during times of no demand, which occurs more often in summer. Naturally, the clustering approach identifies such a pattern as differing, but it is an anomaly with a positive effect on the system. Thus, for fault detection, the statistical method is superior in this case because it would correctly detect no fault.

In Fig. 9 b), the user evidently has no domestic hot water supply. A mere quantification of the temperature difference would lead to a false positive classification as a very low cooling case for the summer data. Thus, this study could provide the boundary condition of the general temperature level based on such patterns. Again, this emphasizes also the necessity of constant monitoring.

Figure 9 c) displays a case in which the return temperature frequently exceeds the supply temperature. A chain of circumstances leads to the pattern. The temperature control of the domestic hot water storage is disabled because the temperature sensor of the tank is not mounted correctly. As the pre-defined control program does not match, an alternative (unsuitable) program is selected. The return temperature limitation is very slow. By heat conduction, the resulting continuous volume flow of the hot water on the secondary side constantly warms the return temperature measuring probe, which is close to the heat exchanger in times of primary side stagnation. For system efficiency, this pattern is disadvantageous because of high return temperatures. User comfort is affected as well, as confirmed by the inhabitants. However, evaluating the average return temperature does not identify the fault, as the low return temperature values balance the very high ones. Thus, solely the evaluation of the temperature inversion this study established by using the clustering approach reveals the fault. The depicted temperature signature is referred to as R in Table 1. The investigation of the model region led to the conclusion that all substations on the road show the same chain of faults.

The patterns in Fig. 9 d) and e) firstly appear similar, yet the faults differ strongly. In Fig. 9 d), a visit revealed an uncontrolled heat exchanger for domestic hot water. An uncontrolled volume flow at its maximum constantly passes. However, the space heating facility shows no anomaly. Thus, the visible pattern is the sum of both. Especially in summer, the impact of the domestic hot water heat exchanger dominates.



**Fig. 9.** Special cases of end-user temperature signatures in the model region. (a to f) Distinguish between heating period (black dots) and summer (grey crosses) with an indication of advantageous (blue) and disadvantageous (red) positions of data samples and reference to the impact on the heating system.

In contrast, in Fig. 9 e), the control of both heat exchangers employs the return temperature limitation as an actor, and no other mechanism exists. The valve executing the limitation is very slow due to aging, so the control's effect is considerably small. Thus, the threshold has been increased above 55 °C, assumingly to reestablish user comfort. Especially for times of no or low demand, the cooling of the heat carrier is very low, but the volume flow is not reduced to zero.

A comparable effect leads to the pattern in Fig. 9 f). The data samples show a boundary to the right for both summer and winter data. These are two different settings of the return temperature limitation valves in the heat exchangers for domestic hot water (the required 55 °C) and space heating (approximately 60 °C).

## Discussion

With the objective of maximizing automation and minimizing computational costs, this study first labeled the most frequently occurring behavioral patterns indicating faults in user substations with expert knowledge based on the clustering results. This enabled the development of features for a statistical fault detection process for the identified categories. The combination with the volume flow data provides a numerical value indicating the impact of the occurring fault on the system.

As shown above, the clustering approach aims to identify temperature signatures indicative of faults by assuming that faults lead to anomalous patterns in the recorded data. The clustering approach employs the temperature data only, as related research<sup>6,15,20,23</sup>, and relies on data accumulation instead of the dimensions of the volume flow and time. The principle does not depend on the explicit establishment of an expected behavior to distinguish between normal and anomalous behavior. Instead, the degree of deviation is determined by an evaluation of how often a specific pattern is identified as an anomaly by the clustering method. Encoding the employed data pre-preparation and clustering methods contributes to automation by reducing the manual processing steps and the number of substations to be manually investigated.

With the clustering approach, the features of a high return temperature and low cooling, known to indicate faults<sup>7</sup>, could be confirmed. A third feature, which has been focused on less so far, the inversion of the potential between supply and return temperature, could be identified. This was enabled by the clustering approach, which objectively returns the anomalous patterns regardless of prior knowledge. The third indicator proved beneficial in detecting additional faulty substations in the system. Hence, subsequent research could apply the principle to other model regions to provide additional features.

With statistically accessible labels for fault detection established, the suggested method for the detection of faulty substations works fully automated and with limited reliance on the sequential order of the time-series data. The faults' impact on the heating network could be quantified automatically based on the volume flow readings. The method can potentially be used with limited adjustments for the detection process in other district heating networks. One requirement is sufficient user data recorded for the primary side of supply and return temperatures, as well as the volume flow in an adequate temporal resolution. The suggested threshold values in this study are case-specific. For example, the values of the calculated heat  $Q_{55}$  and  $Q_{ret}$  depend on the considered period of one month and the dimensions of the heat network. The workflow is suitable for a network with data

acquisition on the user level. However, if the measurements of the domestic hot water and the space heating facilities are separated, the process requires modifications, and different temperature signatures will occur. This study suggests executing the statistical fault detection process monthly, as applied in the literature<sup>7</sup>, to enable quick identification and avoid overlooking faults that occur only for certain demand situations.

The relevance of faults on the user's secondary side has been showcased, e.g., in Ref<sup>16,17</sup>. In the model region, only the primary side user data is recorded. However, once the secondary side data become available on a broad scale, these can contribute additional labels for fault detection, providing another relevant area of interest.

Data quality and quantity inflict additional challenges. As displayed above, the model region cannot provide complete datasets of all substations. As stated, only 167 of the 486 provided a comparable amount of data of 90%. This underlines the necessity for high-quality data for the machine learning approaches again. Connections for wireless data transmission can be lost, especially in summer, as foliage interferes. The lost connections lead to more information collected on the heating season, and statements concerning the summer become more complex. Low rates of collected data can lead to misclassifications when detecting faults as statistical impacts grow in relevance.

Different reasons can explain similar patterns of temperature signatures. As shown by the results, the individuality of the facilities in historically grown heat networks poses the major challenge for identifying the cause leading to a specific temperature signature. Therefore, this article addresses the most common temperature signatures and assigns the most common underlying causes. However, expert knowledge should further contribute to providing underlying phenomena and labeling data.

Fault management contributes to a higher resolution of knowledge on thermal energy systems. The results of this study outline the direct benefit implied by successful fault detection for the energy efficiency of heat networks. Additionally, the indirect advantages of increasing data quality for modeling purposes to enhance system efficiency should be highlighted.

## Conclusions

This study's suggested three-step workflow for fault management on the user level within heat networks aims to contribute to the automation of effective fault management methods using data from a model region in northern Germany. The workflow utilizes available primary-side system data (volume flow, supply, and return temperature) of substations on the user level. Clustering analysis first identifies fault-indicating data patterns, which are labeled with expert knowledge to provide statistically accessible features. Secondly, these features are used in an automated statistical detection process to identify defective substations. Thirdly, a numerical quantification of the fault's impact on the heating system is calculated, relying on the thermodynamic relationship of the temperatures and the volume flow.

The workflow clusters heat maps of the substations' temperature signatures, which fuse the information of the volume flow and the time dependency with the temperature data. With expert knowledge, statistically accessible features for fault detection could be generated for common faults: an exceeded return temperature, very low cooling, and inverted temperatures.

Statistical analysis using the features determined with the clustering approach could automatically detect faulty substations within the model region. Disabled return temperature limitation units, defective motoric valves, and faults in the storage control form the most common underlying causes for faulty substation behavior.

This study automatically quantifies the negative impact of a detected fault on the energy system regardless of the substation size by combining the temperature and volume flow data, which can enable the prioritization of fault elimination measures in practical applications. The workflow further outlines thresholds for maintaining relevance.

This study shows that the temperature signature of substations is suitable for identifying data patterns indicating faults decisively impacting the efficiency of heat networks. The results validate that faulty user substations can be detected successfully with minimal manual processing. While the method proved effective, challenges remain in diagnosing faults remotely, emphasizing the need for further studies in diverse model regions to enhance data labeling.

Fault management increases user comfort and system efficiency and can have economic advantages while also improving data quality for machine and deep learning applications. Thus, the suggested method for fault detection contributes to enhancing fault management, improving data-driven modeling, and increasing operational efficiency in heat networks.

## Employment

Ulf Rieck-Blankenburg has been employed at the Stadtwerke Flensburg GmbH, while this study was conducted.

## Data availability

Raw data are not publicly available for the model region dataset to preserve individuals' privacy under the European General Data Protection Regulation. Anonymized datasets generated during the study are available from the corresponding author upon reasonable request and with the consent of the heating system operator.

Received: 12 June 2024; Accepted: 2 December 2024

Published online: 31 December 2024

## References

1. Ćirić, I., *et al.* Intelligent Heat Demand prediction for Advanced District Heat Plant Control. Zdravković, M., Konjović, Z., Trajanović, M. (Eds.) *ICIST 2020 Proceedings*, 198–201 (2020).

2. Kudela, L., Chýlek, R. & Pospíšil, J. Efficient Integration of ML into District Heating Predictive Models. *Energies* **13**, 6381. <https://doi.org/10.3390/en13236381> (2020).
3. Lee, D., Yoon, S. M., Lee, J., Kim, K. & Song, S. H. Applying Deep Learning to the Heat Infeed Planning Problem in a District Heating System. *Energies* **13**, 6641. <https://doi.org/10.3390/en13246641> (2020).
4. Lund, H. et al. 4th Generation District Heating (4GDH) Integrating smart thermal grids into future sustainable energy systems. *Energy* **68**, 1–11. <https://doi.org/10.1016/j.energy.2014.02.089> (2014).
5. Buffa, S., Fouladfar, M. H., Franchini, G., Lozano Gabarre, I. & Andrés Chicote, M. Advanced Control and Fault Detection Strategies for District Heating and Cooling Systems—A Review. *Appl. Sci.* **11**, 455. <https://doi.org/10.3390/app11010455> (2021).
6. Gadd, H. & Werner, S. Achieving low return temperatures from district heating substations. *Applied Energy* **136**, 59–67. <https://doi.org/10.1016/j.apenergy.2014.09.022> (2014).
7. Experiences from Swedish utilities. Månsson, S., Johansson Kallioniemi, P.-O., Thern, M., Van Oevelen, T. & Sernhed, K. Faults in district heating customer installations and ways to approach them. *Energy* **180**, 163–174. <https://doi.org/10.1016/j.energy.2019.04.220> (2019).
8. Gadd, H. & Werner, S. Fault detection in district heating substations. *Applied Energy* **157**, 51–59. <https://doi.org/10.1016/j.apenergy.2015.07.061> (2015).
9. Månsson, S., Johansson Kallioniemi, P.-O., Sernhed, K. & Thern, M. A machine learning approach to fault detection in district heating substations. *Energy Procedia* **149**, 226–235 (2018). <https://doi.org/10.1016/j.egypro.2018.08.187>
10. Calikus, E., Nowaczyk, S., Sant'Anna, A., Gadd, H. & Werner, S. A data-driven approach for discovering heat load patterns in district heating. *Applied Energy* **252**, 113409. <https://doi.org/10.1016/j.apenergy.2019.113409> (2019).
11. Li, M., Deng, W., Xiahou, K., Ji, T. & Wu, Q. A Data-Driven Method for Fault Detection and Isolation of the Integrated Energy-Based District Heating System. *IEEE Access* **8**, 23787–23801. <https://doi.org/10.1109/ACCESS.2020.2970273> (2020).
12. Neumayer, M. et al. Fault and anomaly detection in district heating substations: A survey on methodology and data sets. *Energy* **276**, 127569. <https://doi.org/10.1016/j.energy.2023.127569> (2023).
13. Dai, X. & Gao, Z. From Model Signal to Knowledge: A Data-Driven Perspective of Fault Detection and Diagnosis. *IEEE Transactions on Industrial Informatics* **9**(4), 2226–2238. <https://doi.org/10.1109/TII.2013.2243743> (2013).
14. Zhao, Y., Li, T., Zhang, X. & Zhang, C. Artificial intelligence-based fault detection and diagnosis methods for building energy systems: Advantages challenges and the future. *Renewable and Sustainable Energy Reviews* **109**, 85–101. <https://doi.org/10.1016/j.rser.2019.04.021> (2019).
15. Sandin, F., Gustafsson, J., Delsing, J. & Eklund, R. Basic methods for automated fault detection and energy data validation in existing district heating systems. *13th International Symposium on District Heating and Cooling 3rd of September - 4th of September Copenhagen Denmark* (2012). <https://urn.kb.se/resolve?urn=urn:nbn:se:ltu:diva-40519>
16. Kramer, W. et al. AI for Failure Detection and Forecasting of Heat Production and Demand in District Heating Networks, IEA DHC Report, (2023).
17. Fabre, A. et al. Cost benefit study on the building secondary network for improving DH performance. IEA DHC/CHP Report, (2023).
18. Gianniou, P., Liu, X., Heller, A., Nielsen, P. S. & Rode, C. Clustering-based analysis for residential district heating data. *Energy Conversion and Management* **165**, 840–850. <https://doi.org/10.1016/j.enconman.2018.03.015> (2018).
19. Månsson, S. et al. A taxonomy for labeling deviations in district heating customer data. *Smart Energy* **2**, 100020. <https://doi.org/10.1016/j.segy.2021.100020> (2021).
20. Calikus, E., Nowaczyk, S., Sant'Anna, A. & Byttner, S. Ranking Abnormal Substations by Power Signature Dispersion. *Energy Procedia* **149**, 345–353 (2018). <https://doi.org/10.1016/j.egypro.2018.08.198>
21. Bahawan, H. et al. Detection and identification of faults in a District Heating Network. *Energy Conversion and Management* **266**, 115837. <https://doi.org/10.1016/j.enconman.2022.115837> (2022).
22. Worden, K. & Dulieu-Barton, J. M. An overview of intelligent fault detection in systems and structures. *Struct Health Monit* **3**(1), 85–98. <https://doi.org/10.1177/1475921704041866> (2004).
23. Zhang, F. & Fleyeh, H. Anomaly Detection of Heat Energy Usage in District Heating Substations Using LSTM based Variational Autoencoder Combined with Physical Model. *15th IEEE Conference on Industrial Electronics and Applications (ICIEA) Kristiansand Norway*, 153–158 (2020). <https://doi.org/10.1109/ICIEA48937.2020.9248108>
24. Balali, Y., Chong, A., Busch, A. & O'Keefe, S. Energy modelling and control of building heating and cooling systems with data-driven and hybrid models—A review. *Renewable and Sustainable Energy Reviews* **183**, 113496. <https://doi.org/10.1016/j.rser.2023.113496> (2023).
25. Dang, L. M. et al. Daily and seasonal heat usage patterns analysis in heat networks. *Sci Rep* **12**, 9165. <https://doi.org/10.1038/s41598-022-13030-6> (2022).
26. Zhang, Z., Chong, A., Pan, Y., Zhang, C. & Lam, K. P. Whole building energy model for HVAC optimal control: A practical framework based on deep reinforcement learning. *Energy and Buildings* **199**, 472–490. <https://doi.org/10.1016/j.enbuild.2019.07.029> (2019).
27. Whang, S.E. & Lee, J.-G. Data collection and quality challenges for deep learning. *Proc. VLDB Endow.* **13** **12**, 3429–3432 (2020). <https://doi.org/10.14778/3415478.3415562>
28. Ntakolia, C., Anagnostis, A., Moustakidis, S. & Karcianas, N. ML applied on the district heating and cooling sector: a review. *Energy Syst* **13**, 1–30. <https://doi.org/10.1007/s12667-020-00405-9> (2022).
29. Pirouti, M., Bagdanavicius, A., Ekanayake, J., Wu, J. & Jenkins, N. Energy consumption and economic analyses of a district heating network. *Energy* **57**, 149–159. <https://doi.org/10.1016/j.energy.2013.01.065> (2013).
30. Yu, Z., Fung, B. & Haghighat, F. Extracting knowledge from building-related data — A data mining framework. *Build. Simul.* **6**, 207–222. <https://doi.org/10.1007/s12273-013-0117-8> (2013).
31. MathWorks: k means - k-means clustering - MATLAB- MathWorks Deutschland. <https://de.mathworks.com/help/stats/kmeans.html> (accessed 11 October 2024).

## Acknowledgements

This research was funded by the Stadtwerke Flensburg GmbH and the Society for Energy and Climate Protection Schleswig-Holstein (Gesellschaft für Energie und Klimaschutz Schleswig-Holstein GmbH, EKSH), grant numbers 8/12-57 and 14/12-34. We would further like to express our gratitude to Stadtwerke Flensburg GmbH for providing the research data. We acknowledge financial support by Land Schleswig-Holstein within the funding program Open Access Publikationsfonds.

## Author contributions

S.A.W. carried out the conceptualization, formal analysis, and investigation, and prepared the original draft. S.A.W. and U.R.-B. developed the methodology. M.F. contributed to the formal analysis and investigation. D.V. acquired the funding with the support of S.A.W., M.F. and D.V. were in charge of supervision. All authors reviewed the manuscript. All authors have read and agreed to the published version of the manuscript.

www.nature.com/scientificreports/

**Funding**

Open Access funding enabled and organized by Projekt DEAL.

**Competing interests**

The authors declare no competing interests.

**Additional information**

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-82103-5>.

**Correspondence** and requests for materials should be addressed to S.A.W.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024

### 3.1.3 Supplementary Material for Publication 1

Supplementary Information

## Clustering- and statistic-based approach for detection and impact evaluation of faults in end-user substations of thermal energy systems

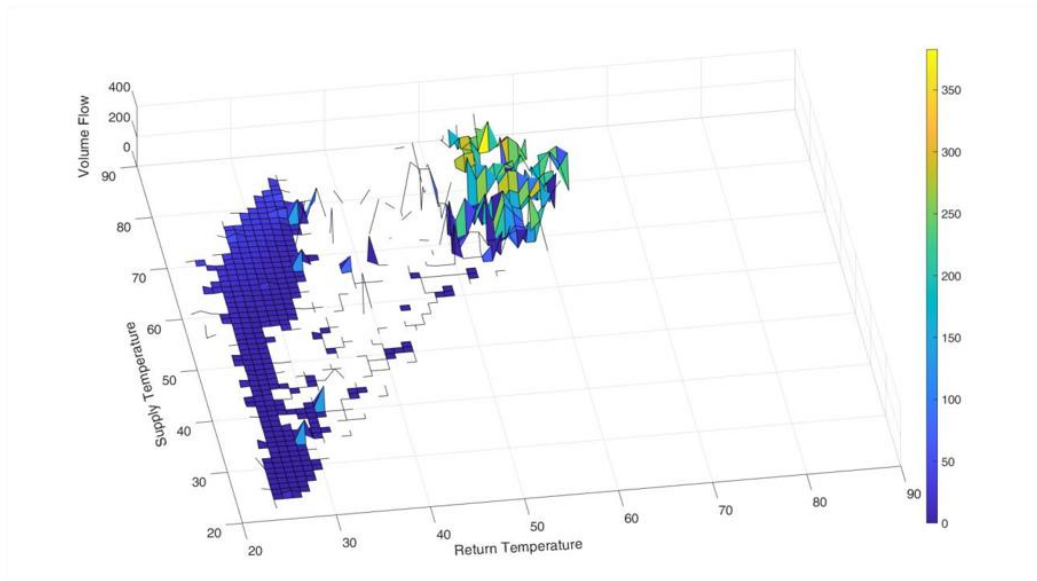
Samanta A. Weber <sup>1,2\*</sup>, Michael Fischlschweiger <sup>2</sup>, Dirk Volta <sup>1</sup> & Ulf Rieck-Blankenburg <sup>3</sup>

<sup>1</sup> Energy and Life Science, University of Applied Sciences Flensburg, 24943 Flensburg, Germany

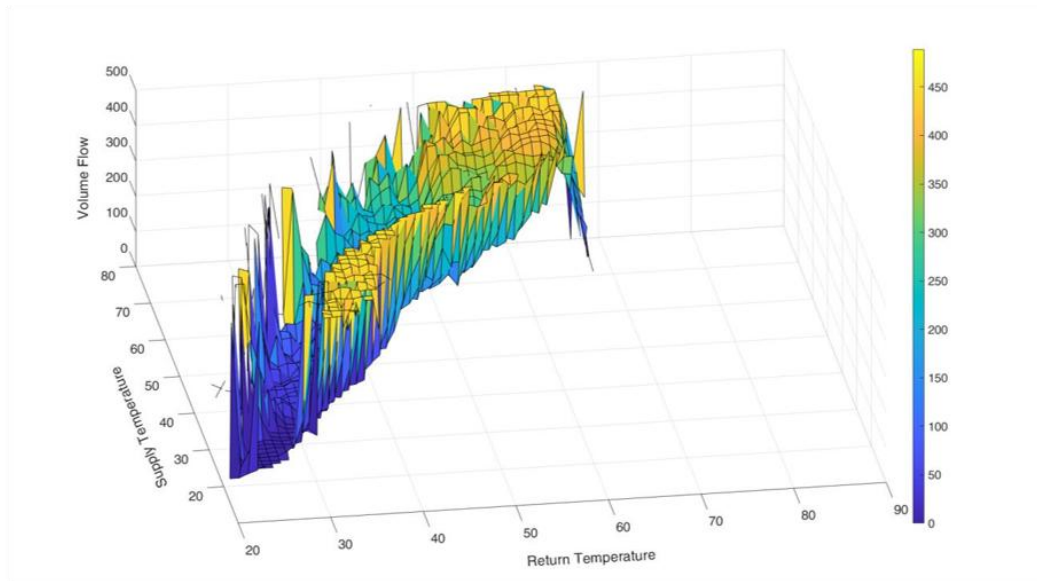
<sup>2</sup> Chair of Technical Thermodynamics and Energy Efficient Material Treatment, Institute of Energy Process Engineering and Fuel Technology, Clausthal University of Technology, 38678 Clausthal-Zellerfeld, Germany

<sup>3</sup> Business Segment Networks, Stadtwerke Flensburg GmbH, 24939 Flensburg, Germany

\*samanta.weber@hs-flensburg.de, samanta.weber@tu-clausthal.de,  
<https://orcid.org/0000-0002-6487-8248>



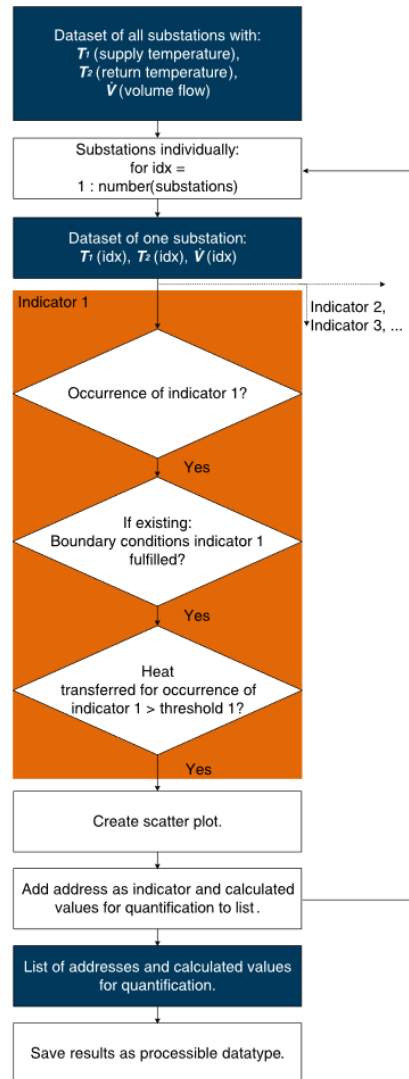
**Figure S1** Volume flow data over supply and return temperature for data of substation in Fig. 4 a), b). High volume flow values occur for a high supply temperature associated with a high return temperature, as the high flow rate leads to reduced cooling of the heat carrier. Low volume flow values correlate with efficient cooling, which can be found on the left.



**Figure S2** Volume flow data over supply and return temperature for data of substation in Fig. 4 c), d). The volume flow values are higher in general compared to Fig. S1. Consequently, the cooling is lower, and the data accumulates more on the right at higher values of the return temperature.



**Figure S3** a) Visualization of processing steps for data pre-preparation and clustering with b) flow chart. The processing steps in orange and the data in dark blue in a) are mirrored by the flow chart in b).



**Figure S4** Flow chart for encoding of statistical fault detection and impact quantification.

For each substation, each indicator for the occurrence of a fault is tested for fulfillment; if existing, indicator-specific boundary conditions are checked; the heat transferred for the occurrence of the indicator is calculated; if the amount is higher than the assigned threshold, the substation's address and computed values are returned.

### **3.2 Publication 2: Feature Selection for Specific Prediction Targets at the User Level in a District Heating Network**

The subobjective of the second publication is to identify a set of relevant, non-redundant features for modeling the supply and return temperatures, as well as the volume flow at the user level of the model region's heat network. Using the suggested parameters as modeling targets within a data-driven modeling framework has been investigated to a limited extent due to data scarcity. Hence, knowledge of the sophisticated inputs to such models is sparse. Consequently, feature engineering and feature selection require investigation.

To achieve subobjective 2, the publication proposes a chain of methods. First, domain knowledge and a thorough literature review are employed to identify a subset of potential features based on commonly used influencing factors in ML-based heat network modeling. Following these steps, statistical and correlation analyses identify features with high information content and detect feature redundancies, which should be avoided. Using the information from the previous steps, feature investigation and selection methods are applied. Based on these, a set of relevant, non-redundant features is identified for the modeling targets.

Achieving subobjective 2, i.e., providing suitable features, is one essential step toward modeling the suggested user-level parameters in heat networks using ML. It is found that, although the features are transferable from modeling other heat network parameters to some extent, the specific relevance of the influencing factors to the modeling target varies. Identifying adequate inputs for an ML model accelerates the model-building process. In the future, the exploitation of the modeled user-level information can enable efficiency and flexibility gains in heat networks. Essentially, within the research design of this cumulative dissertation, the findings gained by pursuing subobjective 2 provide the necessary information for selecting suitable features for subobjective 3 of this research.

#### **3.2.1 Author Contribution Statement of Publication 2**

S.W. carried out the conceptualization, the formal analysis and investigation, developed the methodology and prepared the original draft. M.F. and J.G. contributed to the formal analysis and investigation. D.V. acquired the funding with support of S.W., M.F. and D.V. were in charge of supervision. All authors reviewed the manuscript. All authors have read and agreed to the published version of the manuscript.

#### **3.2.2 Publication 2**

WEBER, S.A./FISCHLSCHWEIGER, M./VOLTA, D./GEISLER, J.: Feature selection for specific prediction targets at the user level in a district heating network. *Sci. Rep.*, 15, 29789, <https://doi.org/10.1038/s41598-025-15777-0>, 2025.

In publication 2, the referencing of sections through consecutive numbering occurs as follows.

## 1 Introduction

## 2 Material and Methods

### 2.1 Model region

### 2.2 Suggested prediction targets and influencing factor determination

### 2.3 Feature engineering for the influencing factors

### 2.4 Statistical and correlation analysis for evaluating feature interdependencies

### 2.5 Feature selection methods

#### 2.5.1 Filter type and embedded type feature selection

#### 2.5.2 Wrapper type feature selection

## 3 Results and Discussion

### 3.1 Statistical properties' visualization for the prediction targets

### 3.2 Statistical properties' exemplary visualization for the meteorological data

### 3.3 Correlation analysis of the features

### 3.4 Feature evaluation with MRMR, F-tests, predictor importance RT, NCA

#### 3.4.1 Volume flow as prediction target

#### 3.4.2 Supply temperature as prediction target

#### 3.4.3 Return temperature as prediction target

#### 3.4.4 Feature selection with MRMR, F-tests, predictor importance RT, and NCA

### 3.5 Feature selection applying RNN for considering temporal dynamic influences

### 3.6 Summary and discussion

## 4 Conclusions



OPEN

# Feature selection for specific prediction targets at the user level in a district heating network

Samanta A. Weber<sup>1,2✉</sup>, Michael Fischlschweiger<sup>1</sup>, Dirk Volta<sup>2</sup> & Jens Geisler<sup>2</sup>

With the challenge of district heating network transition as part of the global objective of clean energy, machine learning provides a methodological approach for understanding the relationships between various influencing factors and demand-side properties of district heating networks, which is decisive for reducing losses, enhancing sustainability, and guaranteeing residential comfort. This work focuses on accelerating the application of modern machine learning methods to modeling district heating networks by generating knowledge on feature engineering and selection for newly suggested prediction targets, namely volume flow, supply, and return temperatures, directly at the building level. A systematic workflow for data acquisition, feature engineering, and selecting the most relevant predictors is presented. For this, statistical and machine learning methods are applied to engineer respective features and establish specific interdependencies, including meteorological conditions, human behavioral patterns, and operational parameters, based on a model region in northern Germany. The qualitative results indicate that the highest impact is for temporal predictors and operational features derived from the infeed facility's data, i.e., approximately 15 to 20% of the total predictor relevance. In comparison to studies targeting the heat load and suggesting outside air temperature as the most relevant predictor, it was found that for the herein proposed prediction targets, this feature is of secondary relevance (roughly 6–10%). The findings of this study provide a feature engineering and selection strategy, as well as relevant knowledge gain, which is a prerequisite for efficient modeling of district heating networks based on machine learning in the future.

Climate change-related phenomena have underscored the need for practical efforts in decarbonizing the energy markets. There should be increased interest in district heating networks, which are considered the most efficient heat supply strategy in urban areas<sup>1–3</sup>. In various engineering fields<sup>4–7</sup>, machine learning (ML) modeling strategies have been developed recently, offering distinct advantages in modeling complex problems and networks. Specifically, for district heating networks, the application of ML approaches could lead to increased automation<sup>8</sup> and contribute to reaching the goals of carbon-neutral, sustainable, and efficient operation of district heating networks (e.g., Ref<sup>2,9,10</sup>). Yet, the information on the sophisticated choice of inputs, especially for modeling user-level parameters in district heating networks using ML, remains limited. Existing work focuses on predicting the heat load at the infeed facility<sup>1,3,11–23</sup>. A limited number of studies investigate the supply<sup>24,25</sup> or return<sup>26,27</sup> temperature as a prediction target. Only recently has the heat demand at user-level substations been investigated as a prediction target<sup>28</sup>. At the same time, findings emphasize the importance of integrating user data into the district heating networks' control mechanisms to achieve efficiency gains<sup>24,26,28–31</sup>. With the potential for applying ML that the existing literature suggests, future research should provide the required knowledge, i.e., information on relevant features, for developing sophisticated ML models in district heating networks for user-level parameters.

The increasing effort to integrate ML methods into energy system operations highlights the relevance of research in this field. The selection of optimized inputs for an ML model is specific to the dataset. As a result, the process of feature selection is an essential part of building an ML model. The literature emphasizes the importance of selecting the most suitable predictors<sup>13,32</sup>. Furthermore, the relevance of selecting appropriate methods for feature extraction is stressed<sup>13</sup>. Inquiring about the most sophisticated feature selection process, existing work highlights the longevity and intense computational costs associated with feature selection, underscoring the benefit of access to domain-related results<sup>11</sup>. The proposition of additional features, which increase predictive information, is named as a relevant novelty<sup>33</sup>. However, in the case of district heating networks, the number of

<sup>1</sup>Chair of Technical Thermodynamics and Energy Efficient Material Treatment, Institute for Energy Process Engineering and Fuel Technology, Clausthal University of Technology, 38678 Clausthal- Zellerfeld, Germany.

<sup>2</sup>Energy and Life Science, University of Applied Sciences Flensburg, 24943 Flensburg, Germany. ✉email: samanta.weber@tu-clausthal.de; samanta.weber@hs-flensburg.de

investigated predictors still tends to be limited, and further research is needed to increase the understanding of influencing factors<sup>15,17</sup>, especially for user-level parameter modeling.

This work focuses on accelerating the application of modern ML methods by generating knowledge on feature engineering and selection to investigate district heating networks for newly suggested user-level prediction targets, namely volume flow, supply, and return temperatures. The selection of the optimal set of inputs for ML models remains dataset-specific, with existing studies highlighting the relevance of feature selection. Yet, literature on the most sophisticated choice of inputs for ML user-level modeling in district heating networks remains sparse. Thus, this study proposes a systematic workflow for domain-specific feature engineering and selection, enabling the quantitative investigation of the relevance of influencing factors and derived features in novel prediction targets. This work utilizes experimental data from an existing district heating network in Germany as a model region for analyzing the interdependencies between operational, temporal, and meteorological parameters and the suggested prediction targets. Hence, the results are presented qualitatively, as they remain specific to this study's model region until proven otherwise. This work relies on statistical and ML-based feature investigation and selection methods. As given by the well-established thermodynamic relationship for heat flow, which equals the product of mass density, specific heat capacity, volume flow, and temperature difference, the selected prediction targets convey thermodynamic information on the district heating network's efficiency. Thus, understanding the output parameters at the user level increases knowledge for optimizing district heating network operation. Modeling user-level parameters with ML methods, therefore, can hold potential for enabling more effective control of district heating networks, reducing heat losses, optimizing load distribution, and facilitating the integration of renewable energy sources—ultimately supporting the transition toward more sustainable and carbon-neutral heating systems.

The article is structured as follows: Sect. 2 outlines the research methodology and input parameters, providing an overview of commonly selected features for heat load prediction. Domain knowledge and the literature provide the potential features for the suggested prediction targets. Further, the section introduces the model region and addresses the applied feature engineering and selection methods. Section 3 presents the case study and its results, which are discussed consequently. Section 4 finishes with the conclusions.

Materials and methods

The suggested systematic feature engineering and selection approach is sequential: first, domain knowledge and literature research are used to identify the factors influencing district heating network parameters in the model region. The acquired raw datasets are then transformed into processable features through feature engineering. These form the set of potential features. Subsequently, the potential feature data are analyzed using statistical property visualization and correlation investigation to gain insight into the information content of the datasets and their interdependencies, thereby identifying redundant features. Consequently, feature evaluation and selection methods, both with and without the ability to utilize information from the sequential order of time series data, are applied to rank the relevance of potential features for the respective prediction targets. Lastly, by combining the information from the preceding steps, the results are interpreted to obtain a relevant and non-redundant subset of features for application in ML modeling.

Model region

Feature selection begins with identifying the factors that influence the prediction targets. These factors are specific to the prediction targets' datasets and, hence, particular to the model region presented in the following. The higher the comparability of a different model region to the one of this study, the higher the transferability of the results will be.

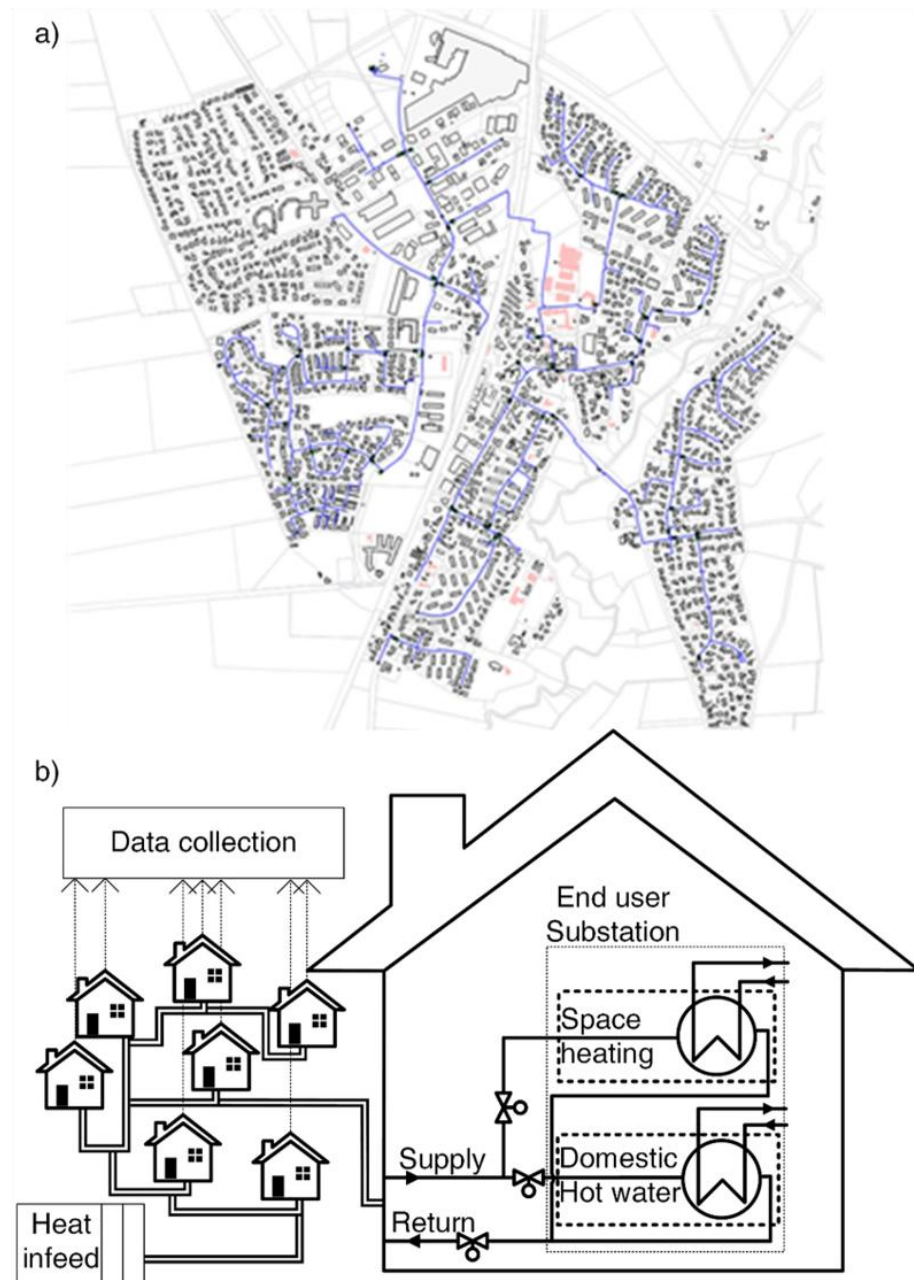
Table 1 presents the technical parameters of the district heating network located in Tarp, Schleswig-Holstein, Germany, which is connected to a combined heat and power plant (CHP). Figure 1a displays the piping structure of the model region, and Fig. 1b schematically presents the network and heat transfer concept in the substations.

The supplied municipality, presented in Fig. 1a, consists mainly of stub lines, with one circular line also existing. As shown in the schematic representation in Fig. 1b, users in the model region can receive heat for both space heating and domestic hot water, resulting in year-round demand. Furthermore, domestic hot water can be supplied using a storage facility or directly, without the need for a storage tank. Deviating from Fig. 1b, it is also possible that the user receives heat solely for space heating.

Of the available network data, the recordings of the suggested prediction targets are collected directly at the substations. The recording of the substations' parameters works remotely. The measuring unit sends momentary values approximately every hour to data accumulators in the network, making them digitally accessible. The address and the timestamp are identifiers. This study utilizes data on volume flow, supply, and return temperatures

| Technical parameter            |                      |
|--------------------------------|----------------------|
| Annual heat supplied           | 26 GWh               |
| Peak load                      | 12 MW                |
| Operational supply temperature | 70 to 105 °C         |
| Return temperature             | 50 to 55 °C          |
| Piping type                    | Plastic sheath pipes |
| Number of substations supplied | 475                  |

Table 1. Technical parameters of the district heating network in the model region.



**Fig. 1.** (a) District heating network piping structure, and (b) schematic heat transfer concept in the model region. The model region comprises a small, centrally supplied district heating network with mainly stub lines and one closed circle line as shown in (a). The figure (a) is reproduced with permission from Stadtwerke Flensburg GmbH. The users can receive both heat for domestic hot water and space heating, as presented in the schematic in (b).

collected at approximately hourly resolution from substations from December 1, 2021, to June 17, 2023. For all substations, only 90% of the data is used for training and, where used, validation. Hence, the most recent 10% of the data is not used in any training or validation step. These are independent and remain to test the performance of a later model.

### Suggested prediction targets and influencing factor determination

For ML modeling, the influencing factors of the prediction targets must be converted into features, i.e., data that is processible by an ML algorithm. Thus, the starting point is to gain knowledge of the prediction targets. To better understand these, it is necessary to know that three basic operational modes can exist for substations, which can be distinguished by the volume flow recorded for each instance. The first operational mode occurs when there is no demand. Either the flow rate is very low for direct domestic hot water supply to maintain the supply temperature, or stagnation can occur in cases of only space heating supply or space heating and domestic hot water supply with a storage facility. If the system stagnates, the heat carrier in the connection line cools down in a non-stationary process. This volume must be replaced when the facility restarts. Hence, restarting is the second operational mode. The volume flow increases significantly, often reaching its maximum, as the supply temperature rises until it meets the required level, typically resulting in a high return temperature. Thirdly, the operation can achieve a relatively stable situation of a continuous, relatively stationary heat supply. The volume flow is adjusted to intermediate values. Visualization of statistical properties for the prediction targets, such as histograms and combination scatter plots, provides insight into their distributions and interdependencies. This information further contributes to determining the relevant influencing factors.

When a model for the substations' parameters is established, the features must provide sufficient information to account for these operational modes. Thorough literature research provides the factors that influence heat load parameters. Regarding the dependency between the heat demand of individual buildings and the resulting overall heat load of the network, we assume a similarity between the respective influencing factors. Thus, Table 2 lists the predictors for district heating network modeling, providing an exemplary reference to related studies.

This study aims to investigate the applicability of these influencing factors in the context of the new prediction targets. Domain knowledge provides insight into the applicability of the referenced influencing factors to the newly suggested prediction targets. Of the parameters listed in Table 2, this study employs the dynamic parameters only, apart from the coordinates, because ML algorithms, as used in this study, cannot derive the same information from static parameters as a physical model. Furthermore, this work does not introduce a linearly increasing feature to account for measurement duration, as it is expected to have low relevance for the suggested prediction targets. Having identified the influencing factors, these must be transferred into input datasets that are processible by an ML algorithm ("feature engineering").

| Feature                               | Impact                                                                                                                                                            | References                 |
|---------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------|
| Factors non-influenceable by operator |                                                                                                                                                                   |                            |
| Hour of the day                       | Use patterns over the day, heat capacities, control of heat exchange                                                                                              | 11,17,18,22,28             |
| Day of the week                       | Use patterns due to workdays and holidays                                                                                                                         | 11,13,22,25,28             |
| Day or month of the year              | Seasonal impact, e.g., heating period and summer                                                                                                                  | 11,13,17,18,22,28          |
| Occurrence of a holiday               | Impact on human behavior                                                                                                                                          | 13,17,18                   |
| Duration of measurements              | Long-term effects                                                                                                                                                 | 13                         |
| Outside temperature                   | Heat demand and losses of the buildings, perceived temperature                                                                                                    | 11,17,22–24,28,29,31,33,42 |
| Ground temperature                    | Heat losses of the distribution pipes                                                                                                                             | 42–44                      |
| Wind speed and direction              | Perceived temperature, heat demand and losses of the buildings                                                                                                    | 15,17,23,27,29,31,33,42    |
| Solar radiation, sun                  | Perceived temperature, heat demand and thermal gains of the buildings                                                                                             | 15,23,28,29,33             |
| Precipitation                         | Perceived temperature, thus heat demand and losses of the buildings, where applicable losses of the pipe system (moisture of the soil)                            | 23,31,43                   |
| Air humidity                          | Perceived temperature, thus heat demand                                                                                                                           | 15,17,24,27,31,33          |
| Pipe diameter                         | Surface of heat transfer decisive for losses, resulting capacity flow impacts temperature loss; the heat capacity depends on the mass                             | 23,43                      |
| Pipe length respectively geography    | The longer the distance to the CHP, the higher the temperature loss in general; pipe length determines the surface of the heat transfer again and thus the losses | 23,43                      |
| Insulation                            | The major impact on the thermal resistance should be located in the insulation                                                                                    | 23,43                      |
| Return temperature                    | Necessary volume flow to meet heat demand, losses of the distribution system, efficiency of CHP                                                                   | 14,31,42,44,45a            |
| Factors influenceable by operator     |                                                                                                                                                                   |                            |
| Heat infeed                           | Over- or undersupply, system efficiency, return temperature                                                                                                       | 14,31,42a                  |
| Volume flow                           | Necessary temperatures to meet heat demand, losses of the distribution system, efficiency of CHP                                                                  | 14,31,42,44,45a            |
| Supply temperature                    | Necessary volume flow to meet heat demand, losses of the distribution system, efficiency of CHP                                                                   | 14,31,42,44,45a            |

**Table 2.** Features for parameter prediction in district heating networks as given by literature. <sup>a</sup>Measured at CHP, only.

### Feature engineering for the influencing factors

The following addresses the process of developing representative features for the newly suggested prediction targets from the influencing factors. Certain influencing factors, such as weather parameters and operational data, can be described using measured values. Additionally, individual behavior reflects demand<sup>29</sup>. To express behavioral patterns as numerical inputs processible by ML algorithms, researchers employ values derived from the timestamp<sup>11,13,17,18,22,25,28</sup>.

The collected raw data requires transformation into processable and compressed information. To obtain relevant features, the data must be as clean as possible. Each time series demands individual pre-processing steps, which are adjusted manually. As newly suggested prediction targets are investigated, this study does not derive features by applying intense processing to the data, as performed in the literature (e.g., Ref<sup>13</sup>). Instead, this first approach relies on experimental characteristics.

Starting with the prediction targets, a resampling is conducted to achieve equal distances of hourly values using linear interpolation. Secondly, missing values are filled with a moving mean using a three-hour window to account for dynamics. After manually examining the data, the quality is assumed to be sufficient to omit outlier cleaning.

Subsequently, the operational and meteorological data demand adequate treatment. The data recorded at the CHP originates from internal monitoring and control of the facilities. Of the available data, the supply and return temperatures, the volume flow, the accumulated heat infeed, the supply and return pressures, and, in addition, the outside temperature form potential features. The cleaning of outliers is performed through filling and detection methods, as outlined in Table S1 (Supplementary Information). The differing dynamics of the features result in respective windows for cleaning, threshold value selection, and replacement values for outliers. Manually conducted visualization and testing of optimal choice of method determined the distinct selection of methods for detection, varying between mean, moving mean, and moving median. Table S1 (Supplementary Information) also provides the calculation rules for the latter. After cleaning the parameters of the accumulated heat, this study calculates the heat flow as the difference between consecutive hourly measurements.

Concerning the meteorological data, only the outside temperature is available at the location. Thus, this work utilizes data from the two geographically nearest weather stations, monitored by the German Meteorological Service (Deutscher Wetterdienst) [35] and located in Leck (1) and Schleswig-Holstein (2), Schleswig-Holstein, Germany. Data quality information is available alongside the data. This study relies on the following data: ground temperature (1 m below the surface), minutes of direct solar radiation, wind speed and direction, absolute and relative moisture, as well as precipitation in hourly values. Firstly, this study equalizes the hourly timestamps with linear interpolation. Secondly, this work treats the missing values and outliers, as given in Table S2 (Supplementary Information).

Temporal values result directly from evaluating the timestamp. The occurrence of a holiday is transferred to a binary time series (0 not a holiday, 1 a holiday<sup>17</sup>). As for the coordinates, a uniform time series of the longitude and latitude, respectively, serves as input.

Furthermore, the data require scaling. Firstly, apart from the uniform time series and the occurrence of a holiday, all mean values of the signals are scaled to zero by subtracting the average. To equalize the peak-to-peak value of all signals to one and avoid the influence of magnitude, this work employs min-max normalization. Finally, any missing timestamps in the input or output data must be removed. For each substation, this leaves only the processable information.

### Statistical and correlation analysis for evaluating feature interdependencies

Visualization of statistical properties provides a first impression of the potential predictors' datasets. It can indicate the information content of a feature based on the distribution and interdependencies of features, as shown in combination scatter plots. Features identified to show low information content can be expected to be assigned a low relevance in subsequent feature investigation and selection.

Further, feature redundancy should be addressed. Applying redundant features, i.e., features that provide overlapping or non-additive information to the model, does not increase the information contained in the feature set. Instead, this slows down computation and can lead to overfitting. Therefore, this study applies the standard method of linear correlation<sup>14,15,29</sup> to quantify the linear interdependency of potential features. A numerical value indicating the intensity of the linear correlation between datasets provides an indicator of the difference in information content between potential features. Thus, a correlation investigation gives a means for identifying feature redundancy. Highly correlated features should not be applied in parallel. Instead, only the most relevant feature of the redundant should be selected. In contrast, a correlation between feature and target is indicative of high significance for the prediction. Non-linear mutual information is not captured in this step. This could include, for example, time-shifted occurrences, dynamic events, or non-linear correlations (such as exponential dependencies), among others.

Once the redundancies among the features are identified, the respective relevance of each feature for the prediction target shall be determined using the so-called feature selection methods.

### Feature selection methods

ML modeling architectures have diverse concepts. Therefore, the chosen methods for determining the feature relevance should reflect this. Thus, this study's applied feature selection methods differ, especially in their respective approach to processing time series data, leading to distinct results. Specific ML algorithms enable the computation of information contained in the sequential order of time series data (e.g., recurrent neural networks, RNN). In the literature, the concurrent finding is that RNN structures, mainly the approaches of Long Short-Term Memory (LSTM) and deep neural networks, are most suitable for modeling thermal networks<sup>1</sup>.

Three general predefined approaches exist, namely filter, embedded, and wrapper type methods, for evaluating relevance and selecting representative features [36]. These are briefly explained below. Furthermore, the respective motivation for the selection of methods used in this study is provided. The feature relevance analysis will be conducted individually for each substation, and the results will then be aggregated (i.e., averaged) across all 475 substations to identify general trends. This approach was chosen to account for the heterogeneity of substation behavior while enabling conclusions regarding the features across the network. The numerical results for predictor relevance will be presented in descending order and subsequently categorized. When a significant difference occurs between two consecutive features in this order, a new category will be selected, ranging from features with the highest impact to those that are dispensable. This study applies MATLAB<sup>®</sup> as the programming environment.

#### Filter type and embedded type feature selection

“Filter type feature selection” estimates the influence of a feature in terms of its relevance for the output using statistical measures<sup>36</sup>. Subsequently, the algorithm assigns a numeric value as an indicator<sup>36</sup>. For “embedded type feature selection”, the algorithm executes the feature importance estimation and the training of the model in parallel. Such algorithms evaluate the features based on the predictive power gained by the chosen model architecture from the inputs<sup>36</sup>. “Wrapper type feature selection” either encloses (forward) or withdraws (backward) features while determining the variation of performance a model achieves with the altered set of input parameters. Providing a termination condition, the model repetitively runs through training until no further improvement occurs<sup>36</sup>.

The applicability of the existing methods depends on the dataset, and not all existing methods are suitable. From the filter type, this study uses the principles of Maximum Relevance—Minimum Redundancy (MRMR), analysis of variance (ANOVA) with f-statistics (F-tests), and Neighborhood Component Analysis (NCA). MRMR searches for features that show the highest correlation with the target of prediction and, at the same time, the most substantial difference from the other features<sup>37</sup>. Reference<sup>37</sup> presents the mathematical formulation. F-tests works with a hypothesis test, resulting in the p-value, which is smaller for a feature of higher influence. Any p-value smaller than  $\text{eps}(0)=4.9407\times 10^{-324}$  will result in an assignment of infinite feature importance<sup>38</sup>. The explicit mathematical formulation used in this study is presented in, e.g., Ref.<sup>39</sup>. The choice of these two methods relies on the general applicability to regression problems. Furthermore, all features are investigated in combination, and manual inclusion or exclusion is possible. A specific evaluation is provided as a numeric ranking. The NCA algorithm is based on a minimization method using the mean value of the leave-one-out loss as an objective function, as mathematically described in Ref.<sup>40</sup>. This work applies this method because it can capture the influence of correlation among the respective predictors with the leave-one-out approach<sup>40</sup>. The method evaluates the features that decrease the loss the most, while those that do not lower the loss enough are given zero influence<sup>40</sup>.

From the embedded type alternatives, this work applies the predictor importance of an automatically optimized regression tree (RT). This structure has the advantage of low computational costs. Furthermore, related studies<sup>16,20</sup> utilize the RT modeling architecture for district heating networks.

The approaches from the filter and embedded type employed in this study do not consider the sequential order of the time series data as additional information. To reflect recurrent behavior as part of this study’s objective, an approach resting on an RNN is additionally applied. The conceptualization with a wrapper type feature selection method is explained below.

#### Wrapper type feature selection

The wrapper type feature selection algorithm requires a criterion function that provides a numeric value as a metric for the inclusion or exclusion of a feature. As RNNs are stated to be one of the most suitable choices for modeling district heating networks<sup>1</sup>, the aim is to test the integration of a recurrent structure into feature selection. In this study, the criterion function trains an RNN and subsequently evaluates the root mean squared error (RMSE) as a metric.

The chosen model must be well-balanced, meaning it is flexible enough to account for the dynamics and avoid underfitting, while also being sufficiently restricted in variance to prevent overfitting. A shallow RNN shall be applied to limit the computational costs. The hidden layer contains ten neurons. The network employs a layer delay of [1:2] to capture the influence of the preceding hour. The maximum number of 200 training epochs limits the process. Further selections of hyperparameters and settings of the ML model are captured in Table 3.

To determine the RMSE as the metric for feature selection, this study employs time series cross-validation (CV). The number of features corresponds to the number of groups to be specified for creating subsets of approximately the same size. The validation data must be sequential for time series CV and should be

| Hyperparameter/setting      | Selection             |
|-----------------------------|-----------------------|
| Layers                      | Input, hidden, output |
| Units in hidden layer       | 10                    |
| Optimizer                   | Levenberg–Marquardt   |
| Initial damping factor      | 0.001                 |
| Maximum validation failures | 6                     |

**Table 3.** RNN model hyperparameters and settings.

representative of the entire dataset. Literature on district heating networks indicates a significant difference in patterns between the heating period and summer<sup>13,33</sup>. Hence, two validation sequences are generated for each substation dataset originating from different seasons to minimize this influence on the validation results. This is achieved by setting the share of data to be selected for validation to the reciprocal of the number of features. Then, this amount is set aside for opposing seasons, beginning with January 1st and July 1st for the first feature, and then rolling out for the remaining features.

Before conducting any investigation based on the RNN, such as CV, it is essential to rule out model underfitting<sup>41</sup>. Therefore, the modeling architecture employed was investigated in this regard using the test data set that was set aside. To validate whether an underfitting of the model influences the performance, the autocorrelation coefficients of the residuals were investigated. A low autocorrelation of the residuals could be determined for an applied time lag. This indicates that the model captures most of the predictable part of the data. Rather than underfitting, the randomness of the data leads to the observed RMSE. Hence, the modeling architecture and CV appear to be applicable.

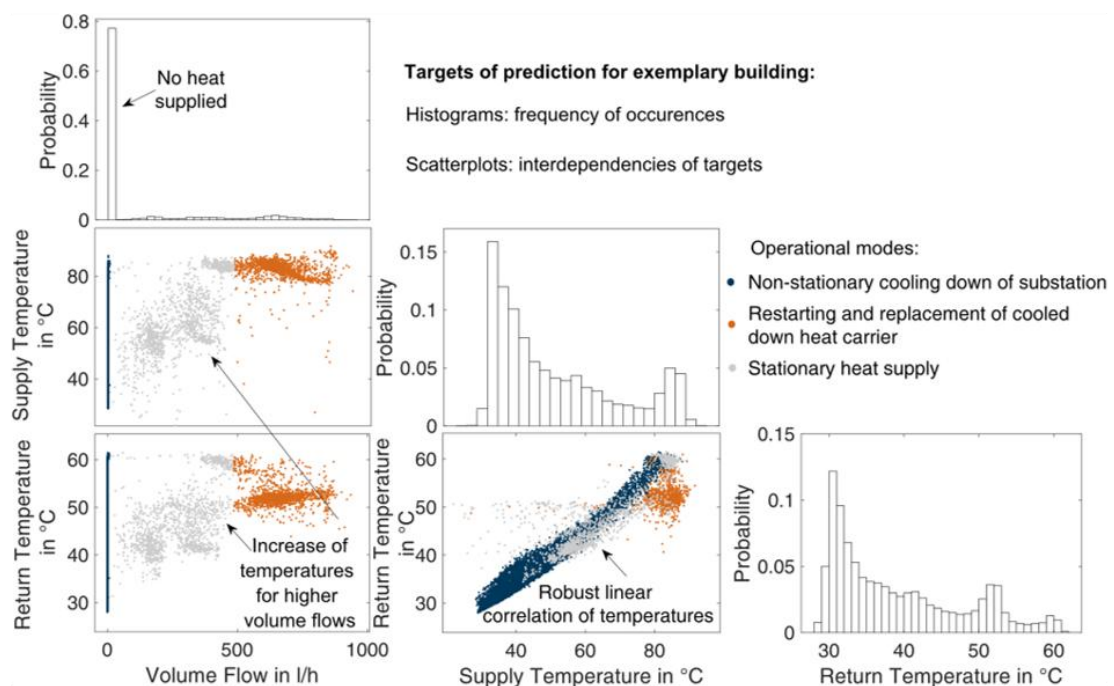
With this setting, the feature is selected if the addition (forward feature selection) to the set reduces the RMSE by at least 0.02. This metric value is experimentally adjusted and specific to our feature set. The overall RMSE differs for the individual buildings but ranges around 0.2. Hence, this study includes a feature if it reduces the RMSE by 10% of the expected value, which is a standard error threshold in significance tests<sup>34</sup>. Unlike the previously mentioned methods, this approach returns a binary list indicating whether a feature is included or excluded, rather than a continuous ranking.

## Results and discussion

The following illustrates the general distributions and interdependencies of the inquired information. First, a substation dataset is visualized to enhance the user's understanding of the data. Second, when starting with the selection of the most relevant potential features, the meteorological data are exemplarily visualized. This highlights how the histograms and combination scatter plots provide an understanding of the information content and the interdependencies of the potential features. Third, linear correlation analysis outlines redundancies among the feature datasets. Lastly, using the gained insights from the previous steps, the feature investigation and selection methods provide a ranking of the quantitative relevance of predictors.

### Statistical properties' visualization for the prediction targets

To deepen the understanding of the prediction targets and the substations' operation, Fig. 2 displays a histogram of the target in the column and a scatter plot of the target in the row against the column for an exemplary



**Fig. 2.** Histograms and combination scatter plots of prediction targets for exemplary building. The histograms of the prediction targets indicate more frequent occurrences. Displaying the data as scatter plots indicates a linear correlation between the temperatures and an increase in temperatures for higher volume flows.

building. The choice of building does not rely on efficient operation, but rather on the suitability of visualizing the three possible operational modes addressed in Sect. 2.2.

To distinguish the operational modes, this work utilizes k-means clustering to separate the volume flow data into three distinct clusters. It applies the restriction that data with zero volume flow indicates stagnation, as the building has a domestic hot water storage facility, viz., the primary-side heat carrier can stagnate. The coloration of the scatter plots in Fig. 2 visualizes the operational modes.

As visible in the data on volume flow, a predominant majority of samples appear at a recording of zero, indicating times when no heat was supplied and limiting the amount of informative data for this target. For the remaining samples, a tendency for higher volume flow is observed, correlating with higher temperatures, indicating the system's restart. Only a minimal amount of data appears for the stationary operation. The distribution of both temperature variables is skewed to the right and bimodal. Moreover, the supply and return temperatures display a robust correlation, particularly for lower values on both temperature scales, highlighting the internal relation.

It should be considered that a later ML model will be less accurate for occurrences with fewer data points, which can necessitate sampling measures. Additionally, specific statistical methods are sensitive to the skewness of distributions. Hence, the investigation also provides information for the intended modeling with ML algorithms.

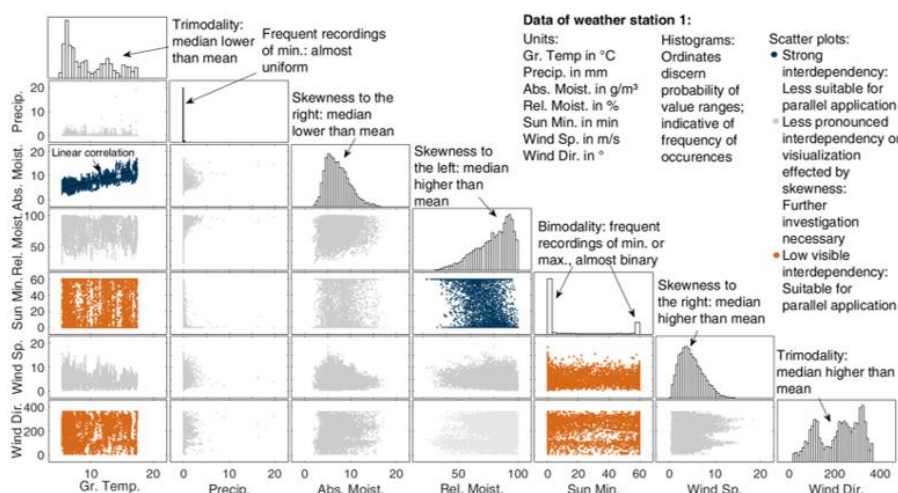
Each substation's data will be unique, resulting from influences such as storage capacities, supply strategies, and user behavior. Thus, the selection of features must contain sufficient information on all the relevant impacts for successful modeling. In the following, the features shall be investigated to determine a subset of relevant, but not redundant, features.

### Statistical properties' exemplary visualization for the meteorological data

Visualization of statistical properties provides an initial impression of the information content of the features and potential redundancies. Exemplarily, Fig. 3 displays the histograms and scatter plots of the parameters for weather station 1. For better visualization, Fig. 3 abandons the titles of the histograms' ordinates, which give the probability of the respective value ranges.

The distributions vary from right-skewed for absolute moisture and wind speed to left-skewed for relative moisture. The precipitation reveals a dominant peak at zero, indicative of high redundancy in the time series and associated low information content. This applies to the bimodal distribution of the sun minutes as well.

Evidently, the ground temperature and absolute moisture exhibit a robust linear correlation, suggesting one of the two is obsolete in modeling. To enable the quantification of existing redundancies, the following presents a correlation analysis of the features.



**Fig. 3.** Histograms and combination scatter plots of parameters for weather station 1. Regarding the distributions of the meteorological data, the skewness indicates more probable ranges of values. Based on the scatter plots, a correlation is evident between ground temperature and absolute moisture, as well as relative moisture and sun minutes, while the other parameters appear to be less strongly related.

### Correlation analysis of the features

Figure 4 displays all possible combinations of two potential features, stating the correlation coefficients, as well as the ranges for evaluating correlation intensity<sup>15</sup> and the resulting evaluation of the suitability of two features to be applied parallelly.

The highest values (dark grey) in Fig. 4 correspond to the correlation of operational data, including supply and return pressure, heat and volume flow, and heat flow and supply temperature. Regarding the weather parameters, the ground temperature and absolute moisture mutually depend on the outside temperature. Furthermore, the operational return temperature correlates with the three named weather parameters.

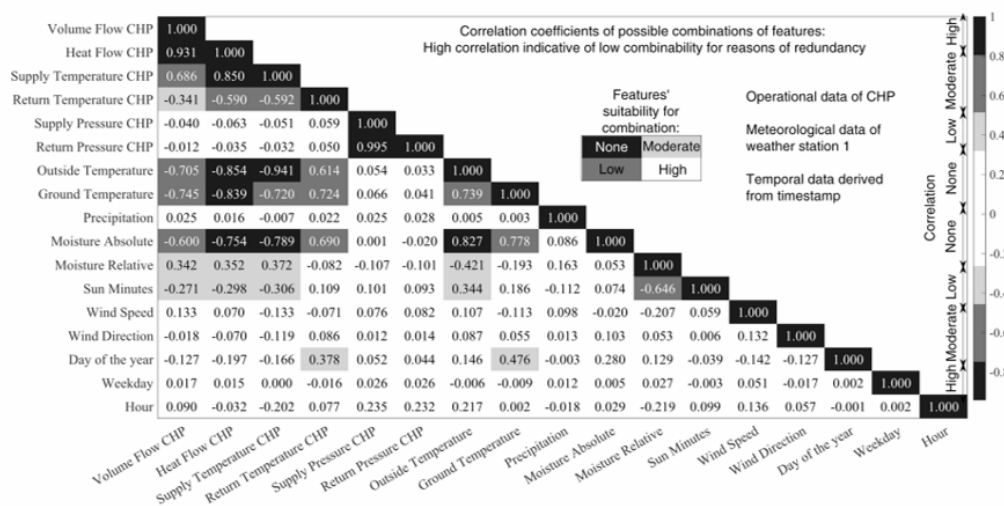
The operational supply temperature, used as a control variable, relies on the outside temperature, resulting in a robust inverse correlation, as indicated by the negative values. Less pronounced, the operational heat and volume flow also exhibit this interdependency. Based on mutual causality, the same applies to the three operational parameters, combined with the ground temperature or absolute moisture. Additionally, the sun minutes and relative moisture appear to be moderately related (lighter grey colors). Furthermore, the combination of the day of the year with the ground temperature and operational return temperature exhibits moderate correlation coefficients, particularly as the return temperature increases towards summer. The remaining factors showing no correlation (white) are potentially independent. Regarding the time series of coordinates (constant values) or the holiday (binary time series), the correlation does not apply, which is why they are excluded from Fig. 4. Since the evaluation is limited to linear correlations, other types of mutual information are not accounted for.

To investigate the applicability of the alternative meteorological data's locations, we conduct a correlation of the two weather stations' recordings, as displayed in Fig. S1 (Supplementary Information). As all parameters, apart from precipitation, correlate moderately to highly, we assume the same for the third location of the model region situated between. Consequently, only the precipitation can potentially not be accounted for. Due to small differences and high computational costs, only the data of weather station 1, located upwind in the main wind direction, is used in the following evaluation. The redundancies among the features are now outlined. However, the relevance of the features for the prediction targets still needs to be determined. Therefore, this study applies feature selection methods to determine the most relevant set of inputs hereafter.

### Feature evaluation with MRMR, F-tests, predictor importance RT, NCA

Firstly, an investigation of the feature set is conducted using the explained filter and embedded type methods. This work interprets the results by employing MRMR, F-tests, predictor importance analysis based on RT, and NCA for the entire dataset. The results are divided by the respective sum to normalize the total of all predictor rankings to 100%. For the F-tests, this study substitutes any infinite values with the maximum value in the respective ranking prior to normalization. Averaging over the individual buildings' results gives the mean value for the respective target of prediction.

The seasonal differences existing in district heating networks<sup>13,33</sup> eventually demand special treatment. To test the influence of such differences on the novel prediction targets, Fig. S2 (Supplementary Information) depicts the results of the separate analysis. For computational efficiency reasons, this study only employs MRMR for this specific investigation. The exhibited discrepancies confirm the need to capture seasonal behavior. Yet,



**Fig. 4.** Evaluation of feature combinability applying linear correlation. The correlation coefficients indicate linear dependencies by very high or low values. If two predictors exhibit a high correlation, they should not be applied parallelly for ML models.

no input parameter would be excluded for one season, which is decisive for the other. If the model architecture cannot account for long-term behavior, it could be more suitable to work with separate predictive models for the summer and the heating period.

Figure 5 Concluding results for MRMR, F-tests, predictor importance RT, and NCA. Regarding the overall results of the predictor evaluation, the general order of relevance is congruent, while the explicit results of the methods differ. Each target of prediction exhibits a respective ranking of relevance.

The specific order of relevance differs for the prediction targets. Hence, the following sections interpret the results for the prediction targets separately. Subsequently, Sect. 3.4.4 outlines the results generalized over all prediction targets.

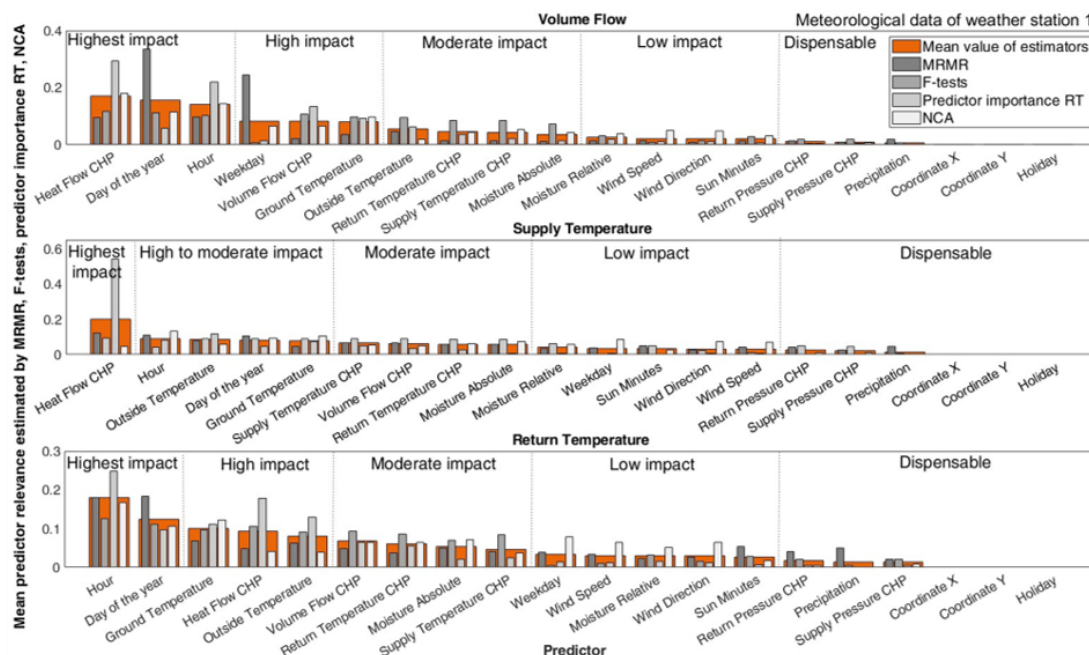
#### Volume flow as prediction target

Starting with the volume flow, the heat flow at the CHP ranks highest with 17.05% of the total feature relevance. Literature confirms the high impact of the operational parameters<sup>31</sup>. Furthermore, the very relevant influence of the temporal values of the hour of the day (15.50%) and the day of the year (14.05%) is evident. Again, the integration of temporal factors by a considerable number of researchers underscores the common conception of their relevance<sup>11,13,17,18,22,25</sup>. For the volume flow at the building substations, the temporal values being among the features with the highest impact suggests that human behavior or control mechanisms based on temporal indicators have a strong influence. In comparison, the weekday (8.21%) ranks considerably lower, indicating less influence.

The algorithms place the volume flow at the CHP next in order, with 8.19%. MRMR and NCA estimate the operational heat and volume flow with a lower influence than the other two algorithms, underlining these features' common causality. This finding concurs with the necessity of eliminating one of these predictors, making the volume flow less relevant.

References<sup>17,18</sup> rank the factors influencing the heat load that they investigated, determining the outside temperature as the most critical factor. However, regarding the ground and outside temperature, following in the ranking with 8.06% and 5.45%, only MRMR would lead to preferring the latter. Concerning the remaining meteorological data, apart from NCA, none of the evaluations would include wind speed (2.08%) or direction (2.05%). Evidently, the operational pressure values, as well as the recordings of the sun minutes and the precipitation, which range between 1.00% and 0.62%, show a very low interdependency with the volume flow at the building.

For the volume flow as a prediction target, the information from the correlation investigation and the feature ranking shall be combined: the outside temperature and absolute moisture, as well as the volume flow at the CHP, should be considered for exclusion due to high correlation. The pressure values at the CHP, alongside the



**Fig. 5.** outlines the final results as the mean of buildings. None of the algorithms reveals any relevance from the coordinates or the holiday time series, which is why they are not discussed further.

sun minutes and the precipitation, could be excluded because of low predictive relevance. The heat flow at CHP and temporal predictors have the highest impact on the volume flow.

#### *Supply temperature as prediction target*

Emphasis should be placed on the findings relating to the supply temperature, as this prediction target has a guaranteed limit not to fall below and is decisive for network losses. By modeling the supply temperature accurately enough, safety margins might be slimmed, resulting in overall more efficient operation.

Remarkably, the most important influencing factor concurrently appears to be the heat flow at the CHP (19.94%), with the RT placing a significantly high impact of 54.20% on it. The hour of the day, outside temperature, day of the year, ground temperature, and operational supply temperature follow in order, accounting for 9.07–7.70% of the total predictor relevance. Although the outside temperature is considered very important by researchers for the heat load as the prediction target<sup>17,18</sup>, other features surpass its relevance for the suggested prediction targets, and the specific model region investigated in this study. Considering Fig. S2 (Supplementary Information), the outside temperature shows a peak for the supply temperature in winter, meaning it is more relevant in the (comparatively longer) heating season.

In descending order, the volume flow at the CHP follows, reaching 6.04%. Yet again, the heat flow at the CHP exhibits a higher influence. In comparison to the results for the volume flow, the weekday ranks much lower (3.28%).

Hence, the ground temperature and absolute moisture, as well as the volume flow at the CHP, seem less suitable for modeling because of their high correlation. The wind speed and direction, as well as the pressure values at the CHP, and the precipitation, also seem insignificant, with values of 3.15–1.29%. The supply temperature relies mainly on the heat flow at the CHP.

#### *Return temperature as prediction target*

Looking at the return temperature, the pattern of the first predictors in the ranking remains roughly the same. Predominantly influenced by the hour of the day (18.05%) and the day of the year (12.44%), the results emphasize the influence of the temporal variables again, even more than for the other two targets, yet placing a significant role on the ground temperature (9.95%) and the heat flow at the CHP (9.28%). MRMR finds more predictive power in the temporal variables, resulting in lower values for the remaining features. Furthermore, this algorithm detects a significant relevance of the precipitation (5.05%). Also, the RT suggests considerable relevance of the hour of the day (24.87%). Concurrently, the ground temperature ranges before the outside temperature and absolute moisture. Only NCA places higher relevance on the weekday (7.86%), as well as the wind speed and direction and the relative moisture (6.50–5.07%), which indicates their low redundancy.

Therefore, the pressure values at the CHP and the precipitation appear obsolete because of their low significance (1.78–1.34%). However, the volume flow at the CHP, the outside temperature, and absolute moisture should be investigated further to determine their relevance because of their high correlation. The return temperature is primarily influenced by the temporal values, and the ground temperature.

#### *Feature selection with MRMR, F-tests, predictor importance RT, and NCA*

The evaluation underlines the highest influence of the temporal features and the heat flow at the CHP. The high relevance of the ground temperature for the chosen model region emphasizes the individuality of feature relevance depending on the prediction target. The results for the suggested prediction targets differ from the results for modeling the heat load, for which the common notion<sup>11,17,22–24,27,29,31,33,42</sup> states a supreme relevance of the outside temperature, which is often the only considered external input<sup>11,25</sup> and claimed<sup>29</sup> or even ranked with feature investigation methods<sup>17,18</sup> to be the most significant factor. At the same time, only a few studies integrate the ground temperature<sup>42–44</sup>. Yet, the findings indicate that for both targets of volume flow and return temperature, the ground temperature holds more predictive power than the outside temperature.

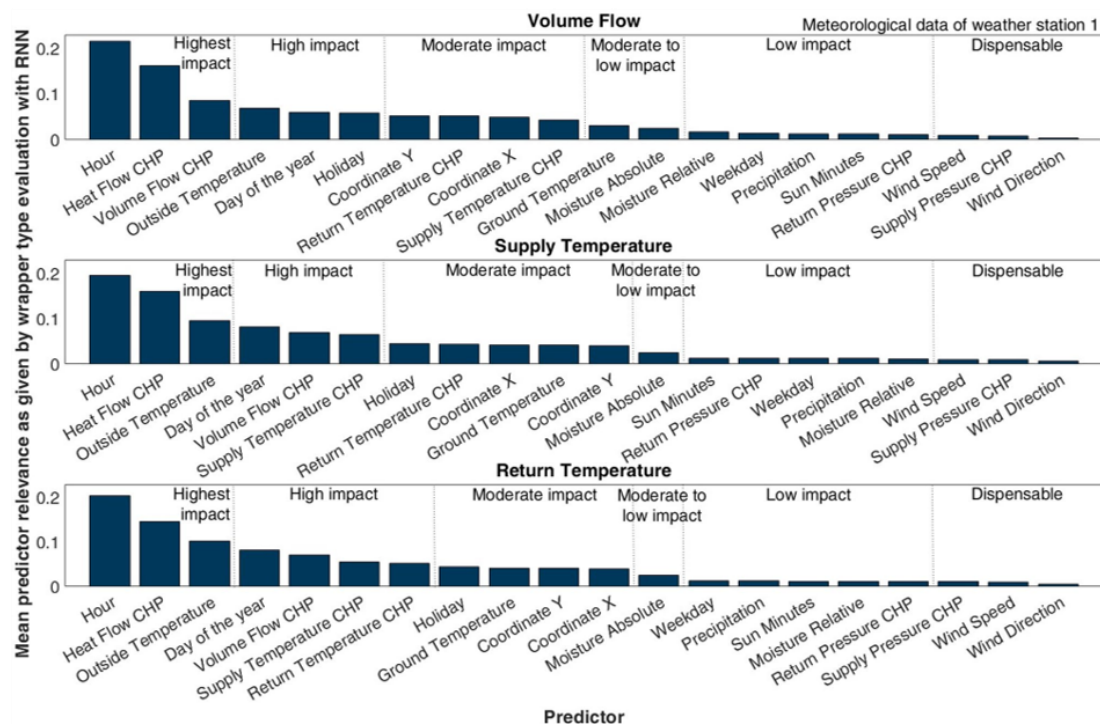
The following compares the results obtained so far to those of the approach using an RNN with wrapper type feature selection to capture the influence of the information contained in the sequential order of the time series data.

#### **Feature selection applying RNN for considering Temporal dynamic influences**

Figure 6 displays the resulting ranking by applying the wrapper-type algorithm in combination with an RNN described in Sect. 2.5.2.

For all prediction targets, the hour of the day (21.61–19.58%) and the heat flow at the CHP (16.24–14.69%) exhibit the highest impact. The remaining features with highest or high impact differ for the prediction targets in descending order. Still, they concurrently include the volume flow at the CHP (8.62–7.05%), the day of the year (8.27–6.06%), and the outside air temperature (10.31–6.87%). The differences in relevance between the prediction targets for the features with moderate to low impact and the dispensable features are minor but result in specific descending orders. Confirming the preceding findings, the operational volume flow is assigned a high relevance, but it is lower than that of the heat flow at the CHP. Additionally, in line with previous findings, the weekday has a lesser impact than the other two temporal values. Concurring with the earlier statements, neither the pressure values at the CHP nor the precipitation significantly influence the prediction targets (1.31–0.55%). However, the wind speed and direction are found to be at the lower end of significance for this evaluation, meaning the RNN can gain even less information from them than the first set of methods. The relative moisture also provides less information in this setting than for the first set of estimations.

Differing from the first investigation with non-recurrent algorithms, the hour of the day clearly achieves the highest rankings, indicating that the recurrent structure detects more influence in this temporal variable. In contrast to the non-recurrent analysis, the outside temperature surpasses the ground temperature (e.g., for



**Fig. 6.** Predictor ranking using RMSE of RNN for criterion function. The approach relies on a recurrent structure, enabling the factoring in of a certain degree of dynamics, which results in a change in the relevance of specific features compared to the former analysis. Yet, the highest impact of temporal variables and the operational data remains valid.

the supply temperature as prediction target 9.55% vs. 4.24%) and clearly holds high relevance. Evidently, the recurrent structure can gain more information from the comparatively more dynamic outside temperature than the set of methods in the first investigation approach. Additionally, the results for the temperatures of the ground and the air differ, which can be explained by the distinct characteristic of the feature selection methods used. As the ground temperature tends to exhibit slower temporal variation, it may capture the baseline thermal conditions affecting buried infrastructure and heat losses in underground distribution pipes more adequately. This could explain its higher relevance in static feature selection methods that prioritize stable, long-term patterns. In contrast, the RNN model may process short-term temporal dynamics more successfully and therefore benefit more from the rapidly changing outside air temperature, which directly influences short-term demand fluctuations, occupant behavior, and heating control responses.

Moreover, the RNN appears to be able to account for the influence of the holiday (5.93–4.47%), indicating that this input should not be neglected when using recurrent structures for modeling. The coordinates are integrated as a uniform time series. An RNN can account for an offset with the bias, but the model still selects these predictors. It is concluded that a uniform time series of ones could be used as an additional feature depending on the choice of modeling architecture. As the other results, in essence, follow the conclusions already drawn, the sequential feature selection confirms the remaining choice of inputs.

### Summary and discussion

In combination with the correlation analysis results, the features are in-/excluded as given in Table 4. A reason for the in- or exclusion is stated for each feature. Certain additional aspects should be considered, as discussed below.

Firstly, the quality and quantity of the data influence the results of our study. The primary objective of applying ML is to identify the general pattern in the data, without relying on noise. Notably, the heat flow at the CHP was found to exhibit higher predictive power than the correlated volume flow at the CHP. The operational heat flow, as described, is calculated using the cumulated heat, a dataset that can be cleaned very effectively, as the value should only increase. Thus, the data on the operational heat flow may contain less noise than that of the volume flow at the CHP, even though both values undergo a cleaning process. In the same context, it is possible that the precipitation recordings do not match the location of the model region. The results must be understood under this restriction. Yet, under the assumption of adequate data collection, the general findings should be

www.nature.com/scientificreports/

| Feature                | Reason for in-/exclusion                                                                                                                                                                                                                   |
|------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Heat flow CHP          | Necessary for closing the loop between operator and building, high correlation with volume flow CHP, chosen for higher rankings                                                                                                            |
| Volume flow CHP        | Comparatively high correlation with heat flow at CHP, but achieves lower rankings                                                                                                                                                          |
| Supply temperature CHP | Necessary for closing the loop between operator and building, but high correlation with heat flow at CHP and outside temperature                                                                                                           |
| Return temperature CHP | Necessary for closing loop between operator and building. Relatively low impact but not dispensable                                                                                                                                        |
| Supply pressure CHP    | Redundant due to almost exact correlation with return pressure, no relevant influence determined                                                                                                                                           |
| Return pressure CHP    | Redundant due to almost exact correlation with supply pressure, no relevant influence determined                                                                                                                                           |
| Outside temperature    | Higher dynamics than ground temperature, decisive for supply temperature at CHP, yet redundancy with ground temperature, importance for supply temperature more pronounced, higher relevance for application of recurrent neural structure |
| Ground temperature     | Decisive for losses of the system, lower dynamics than outside temperature, redundancy with outside temperature, higher relevance for application of statistical model                                                                     |
| Absolute moisture      | Strongly correlated with outside and ground temperature, but not dispensable                                                                                                                                                               |
| Relative moisture      | High correlation with sun minutes, but achieves higher rankings, small influence in general, determined as dispensable using RNN                                                                                                           |
| Sun minutes            | Increased importance in summer, therefore neglectation regarded as critical for this season, despite low rankings overall                                                                                                                  |
| Wind speed             | Low rankings overall, determined as dispensable using RNN, test for recurrent algorithm architectures with long-term memory, as decisive for dynamics of weather phenomena                                                                 |
| Wind direction         | Low rankings overall, determined as dispensable using RNN, test for recurrent algorithm architectures with long-term memory, as decisive for direction of weather phenomena                                                                |
| Precipitation          | Excluded for means of low relevance. Possibly, low rankings due to low applicability to location of model region and highly skewed distribution. MRMR suggests higher relevance, only                                                      |
| Holiday                | Applicable for recurrent networks                                                                                                                                                                                                          |
| Day of the year        | High predictive power, therefore considered decisive for modeling of seasonal impact                                                                                                                                                       |
| Weekday                | Important for users with highly constant patterns (industrial, residential complex), more predictive relevance for summer                                                                                                                  |
| Hour of the day        | High predictive relevance, therefore considered decisive for modeling of human behavior                                                                                                                                                    |
| Coordinates            | Transfer to one uniform time series input                                                                                                                                                                                                  |

**Table 4.** Selection of features based on results of feature evaluation methods.

independent of this circumstance, such as the highest impact of temporal and operational predictors compared to the meteorological data.

Furthermore, the results of this research indicate the relevance of distinguishing between the summer and heating periods. Again, this finding is specific to the location investigated.

Moreover, the list of investigated influencing factors cannot claim completeness because the findings for the newly suggested targets can be cross-referenced only to a limited extent. Based on our work, research can focus on further developing the list of influencing factors and determining the most suitable manner of expressing their impact in numerical terms. It is also possible, given that the targets appear to be related, to consider using one target as input to the other.

Of the selected algorithms for feature investigation, only one relies on a recurrent structure. In this first example, our study employed a simple structure of a shallow RNN. Thus, it would be interesting for future studies to employ such architectures as well.

Aiming to integrate ML models for district heating networks, our work provides a starting point for modeling individual substations. If the black-box architecture and dependencies of the data return a model that is robust enough, the framework could be transferred to other district heating networks. If adequate accuracy can be achieved, an increase in system efficiency may become possible by slimming margins and reducing the need for a physical digital twin model. While ML models offer the potential for automation of the model-building process<sup>8</sup>, they may lack the robustness and interpretability that physics-informed models provide, particularly in safety-critical or extrapolative scenarios [46]. Thus, future research could investigate the most sophisticated level of integrating physical information and automating model building using ML.

Most importantly, until proven otherwise, it must be assumed that the findings of this study remain applicable only to the model region. Naturally, the size of the thermal network, along with its age, maintenance status, connection rate, demand density, and meteorological conditions, as well as social structures, affect the individual results of each district heating network in question. Thus, this study aims not only to provide information on the chosen model region but also to increase knowledge of district heating networks in general by identifying influencing factors and outlining a feature investigation workflow. Future research comparing the explicit findings of this study to those of other model regions would increase the benefit.

## Conclusions

This study aims to accelerate the application of ML in district heating networks by determining the required knowledge on feature engineering and selection for newly suggested user-level prediction targets, namely volume flow, supply, and return temperatures. It employs statistical and ML methods to investigate the feature engineering and selection process at the building level for a district heating network in northern Germany. The study first determines influencing factors and engineers relevant features from the raw data. Second, this work identifies redundancies among the features and assesses the relevance of each feature. The objective is to select a subset of non-redundant, relevant features for ML modeling purposes. This study further compares

statistical and ML-based feature evaluation and selection methods with a feature selection algorithm based on an RNN to assess the impact of incorporating additional information from the dynamic behavior contained in the sequential order of the time series data.

A significant finding is the influence of temporal predictors and operational data from the infeed facility. Notably, specific temporal features, i.e., the time of day and the day of the year, have the highest impact on the suggested prediction targets, accounting for approximately 15–20% of the total predictor relevance. In comparison, the meteorological inputs were identified to have lower relevance for the proposed prediction targets. However, specific meteorological features were identified as having a significant influence on the prediction targets for the model region of this study, including outside air temperature, with roughly 6–10%, and ground temperature, with about 4–10% of the overall predictor relevance. The remaining meteorological features should not be entirely excluded from consideration either, as their relevance may vary across different regions.

This work suggests that the new prediction targets require individual feature engineering and selection processes. Most research combining district heating networks and ML approaches focuses on modeling the heat load. For this prediction target, the outside air temperature is determined as the most relevant predictor. In comparison, the outside air temperature shows less significance for the herein-suggested prediction targets.

Comparing the results in dependence on the application of recurrent or non-recurrent feature evaluation and selection methods, the approach resting on an RNN proves a surplus of information to be gained, e.g., from the higher dynamics of the outside temperature. With the first set of non-recurrent methods, only for the heating period and the target of the supply temperature, the highest impact of the outside temperature could be confirmed. Yet, in this specific case, depending on the applied method, the ground temperature can hold almost as much (around 8% for the supply temperature as the prediction target) or even more predictive power (up to 10%) for the targets of volume flow and return temperature.

The study also reveals that individual building behaviors vary, suggesting that averaging over several buildings can enhance the generalizability of conclusions across the district heating network. The investigation supports the notion that feature engineering plays a significant role in improving the processability of inputs. In this case, creating a feature for the heat flow and generating features from the time stamp proved especially beneficial. It further suggests separating the summer and the heating period in later modeling. An additional feature, structured as a uniform time series, can increase performance, depending on the model.

With a focus on increasing knowledge of energy systems, this study addresses the challenge of engineering and selecting relevant features for modeling user-level parameters in existing district heating networks. The findings suggest that modeling district heating networks requires a feature engineering and selection strategy customized for the specific district heating networks and prediction targets. The study presents a workflow and strategy for ML modeling at the user level in district heating networks. This aims to advance the sustainable and efficient operation of district heating networks as part of the global energy transition.

### Data availability

Raw data for the model region dataset are not publicly available to preserve individuals' privacy under the European General Data Protection Regulation. Anonymized datasets generated for this study are available from the corresponding author upon reasonable request and with the heat network operator's consent.

Received: 12 January 2024; Accepted: 11 August 2025

Published online: 14 August 2025

### References

- Ćirić, I. et al. Intelligent heat demand prediction for advanced district heat plant control. In *ICIST 2020 Proceedings*, 198–201 (2020).
- Kudela, L., Chýlek, R. & Pospíšil, J. Efficient integration of ML into district heating predictive models. *Energies*. **13** (23), 6381. <https://doi.org/10.3390/en13236381> (2020).
- Lee, D., Yoon, S. M., Lee, J., Kim, K. & Song, S. H. Applying deep learning to the heat infeed planning problem in a district heating system. *Energies*. **13** (24), 6641. <https://doi.org/10.3390/en13246641> (2020).
- Moslehi, M. H. Photocatalytic decomposition of metronidazole by zinc hexaferrite coated with bismuth Oxydioxide magnetic nanocomposite: advanced modelling and optimization with artificial neural network. *Chemosphere*. **356**, 141770. <https://doi.org/10.1016/j.chemosphere.2024.141770> (2024).
- Arghavan, F. S. et al. Complete degradation of Tamoxifen using FeNi<sub>3</sub>@SiO<sub>2</sub>@ZnO as a photocatalyst with UV light irradiation: A study on the degradation process and sensitivity analysis using ANN tool. *Mater. Sci. Semiconduct. Process.* **128**, 105725. <https://doi.org/10.1016/j.mssp.2021.105725> (2021).
- Al-Musawi, T. J. et al. Adsorption of malachite green dye onto almond Peel waste: a study focusing on application of the ANN approach for optimization of the effect of environmental parameters. *Biomass Convers. Biorefinery*. **13** (13), 12073–12084. <https://doi.org/10.1007/s13399-021-02174-6> (2022).
- Allahyari, E., Gholami, A., Arab-Zozani, M., Ameri, H. & Nasseh, N. Using emotional intelligence to predict job stress: Artificial neural network and regression models. *DOAJ (DOAJ: Directory Of Open Access Journals)* (2021). <https://doaj.org/article/741d211e29ea41dd97d480b3edd17221>
- Mbiydzennyuy, G. et al. Opportunities for machine learning in district heating. *Appl. Sci.* **11**, 6112. <https://doi.org/10.3390/app1136112> (2021).
- Wang, J., Pan, B., Rong, L., Sun, C. & Qi, C. Topology reconstruction of the district heating network for maximizing comprehensive benefits of thermal storage. *Energy*. **313**, 134065. <https://doi.org/10.1016/j.energy.2024.134065> (2024).
- Wang, J., Jin, Y., Zheng, J., Xue, P. & Sun, C. Optimal quantity, location and capacity allocation of the additional renewable energy stations in a large-scale district heating system and comprehensive case analyses. *Energy. Conv. Manag.* **284**, 116975. <https://doi.org/10.1016/j.enconman.2023.116975> (2023).
- Suryanarayana, G., Lago, J., Geysen, D., Aleksiejuk, P. & Johansson, C. Thermal load forecasting in district heating networks using deep learning and advanced feature selection methods. *Energy*. **157**, 141–149. <https://doi.org/10.1016/j.energy.2018.05.111> (2018).

12. Ali, H., Dedov, A. & Volkova, A. Exploring heat demand forecasting in district heating networks using random parameter linear regression model. *Environ. Clim. Technol.* **28** (1), 670–685. <https://doi.org/10.2478/rtuct-2024-0052> (2024).
13. Potočník, P., Škerl, P. & Govekar, E. Machine-learning-based multi-step heat demand forecasting in a district heating system. *Energy Build.* **233**, 110673. <https://doi.org/10.1016/j.enbuild.2020.110673> (2021).
14. Xue, G. et al. District heating load prediction algorithm based on feature fusion LSTM model. *Energies*. **12** (11), 2122. <https://doi.org/10.3390/en12112122> (2019).
15. Lv, R., Yuan, Z., Lei, B., Zheng, J. & Luo, X. Building thermal load prediction using deep learning method considering time-shifting correlation in feature variables. *J. Build. Eng.* **61**, 105316. <https://doi.org/10.1016/j.job.2022.105316> (2022).
16. Wei, Z. et al. Prediction of residential district heating load based on ML: A case study. *Energy*. **231**, 120950. <https://doi.org/10.1016/j.energy.2021.120950> (2021).
17. Dang, L. M. et al. Daily and seasonal heat usage patterns analysis in heat networks. *Sci. Rep.* **12**, 9165. <https://doi.org/10.1038/s41598-022-13030-6> (2022).
18. Dang, L. M. et al. Toward explainable heat load patterns prediction for district heating. *Sci. Rep.* **13**, 7434. <https://doi.org/10.1038/s41598-023-34146-3> (2023).
19. Calikus, E., Nowaczyk, S., Sant'Anna, A., Gadd, H. & Werner, S. A data-driven approach for discovering heat load patterns in district heating. *Appl. Energy*. **252**, 113409. <https://doi.org/10.1016/j.apenergy.2019.113409> (2019).
20. Finkenrath, M., Faber, T., Behrens, F. & Leiprecht, S. Holistic modelling and optimisation of thermal load forecasting, heat generation and plant dispatch for a district heating network. *Energy*. **250**, 123666. <https://doi.org/10.1016/j.energy.2022.123666> (2022).
21. Golla, A., Geis, J., Loy, T., Staudt, P. & Weinhardt, C. October. An operational strategy for district heating networks: application of data-driven heat load forecasts. *9th DACH + Conference on Energy Informatics*, Sierre, Switzerland, 29–30. *Energy Inform.* **3**(Suppl 1), 1–22 (2020).
22. Johansson, C. et al. Operational demand forecasting in district heating systems using ensembles of online machine learning algorithms. *Energy Proc.* **116**, 208–216 (2017).
23. Lumberras, M. et al. Data-driven model for heat load prediction in Buildings connected to District Heating by using smart heat meters. *Energy*. **239**(Part D) (2022).
24. Zhong, W., Feng, E., Lin, X. & Xie, J. Research on data-driven operation control of secondary loop of district heating system. *Energy*. **239**(Part B), 122061. <https://doi.org/10.1016/j.energy.2021.122061> (2022).
25. Laakkonen, L. et al. Predictive supply temperature optimization of district heating networks using delay distributions. *Energy Proc.* **116**, 297–309. <https://doi.org/10.1016/j.egypro.2017.05.076> (2017).
26. Sun, C. et al. An intelligent control and regulation strategy aiming at Building level heating balance in district heating system. *Energy*. **278**(Part B), 127941. <https://doi.org/10.1016/j.energy.2023.127941> (2023).
27. Gong, M. et al. Gradient boosting machine for predicting return temperature of district heating system: A case study for residential buildings in Tianjin. *J. Build. Eng.* **27**, 100950. <https://doi.org/10.1016/j.job.2019.100950> (2020).
28. Kemper, N., Heider, M., Pietruschka, D. & Hähner, J. Forecasting of residential unit's heat demands: a comparison of machine learning techniques in a real-world case study. *Energy Syst.* **16**, 281–315. <https://doi.org/10.1007/s12667-023-00579-y> (2025).
29. Sun, C. et al. Research on control strategy integrated with characteristics of user's energy-saving behavior of district heating system. *Energy* **245**, 123214. <https://doi.org/10.1016/j.energy.2022.123214> (2022).
30. Li, Z. et al. Novel effective room temperature-based predictive feedback control method for large-scale district heating substation. *Appl. Therm. Eng.* **218**, 119241. <https://doi.org/10.1016/j.applthermaleng.2022.119241> (2023).
31. Wang, Y. et al. Data-driven analysis and prediction of indoor characteristic temperature in district heating systems. *Energy*. **282**, 129023. <https://doi.org/10.1016/j.energy.2023.129023> (2023).
32. Fallah, S. N., Ganjkhani, M., Shamshirband, S. & Chau, K. -w. Computational intelligence on Short-Term load forecasting: A methodological overview. *Energies*. **12** (3), 393. <https://doi.org/10.3390/en12030393> (2019).
33. Kurek, T. et al. Heat demand forecasting algorithm for a Warsaw district heating network. *Energy*. **217**, 119347. <https://doi.org/10.1016/j.energy.2020.119347> (2021).
34. Stevens, J. P. *Applied Multivariate Statistics for the Social Sciences*, 5th ed. (Routledge, 2009).
35. Index of climate\_environment/CDC/observations\_germany/climate/hourly/. [https://opendata.dwd.de/climate\\_environment/CDC/observations\\_germany/climate/hourly/](https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/hourly/) (accessed 14 July 2023).
36. Introduction to Feature Selection—MATLAB, Simulink—MathWorks Deutschland. [https://de.mathworks.com/help/stats/feature-selection.html#mw\\_ef6af785-15ec-4128-a6c2-8d29614825b7](https://de.mathworks.com/help/stats/feature-selection.html#mw_ef6af785-15ec-4128-a6c2-8d29614825b7) (accessed 14 July 2023).
37. Ding C. & Peng H. Minimum redundancy feature selection from microarray gene expression data. *J. Bioinform. Comput. Biol.* **3**(2), 185–205. <https://doi.org/10.1142/s0219720005001004> (2005).
38. Univariate feature ranking for regression using F-tests- MATLAB fsrftest—MathWorks Deutschland. <https://de.mathworks.com/help/stats/fsrftest.html> (accessed 11 Nov 2023).
39. Rasmussen, C. E. et al. *The DELVE Manual*. (MLG Cambridge, 1996).
40. Neighborhood Component Analysis (NCA) Feature Selection—MATLAB, Simulink—MathWorks Deutschland. <https://de.mathworks.com/help/stats/neighborhood-component-analysis.html> (accessed 18 Aug 2023).
41. Hewamalage, H., Ackermann, K. & Bergmeir, C. Forecast evaluation for data scientists: common pitfalls and best practices. *Data Min. Knowl. Discov.* **37**(2), 788–832. <https://doi.org/10.1007/s10618-022-00894-5> (2022).
42. Chicherin, S., Mašatin, V., Siirde, A. & Volkova, A. Method for assessing heat loss in a district heating network with a focus on the state of insulation and actual demand for useful energy. *Energies*. **13**(17), 4505. <https://doi.org/10.3390/en13174505> (2020).
43. Jakubek, D. et al. Mathematical modelling and model validation of the heat losses in district heating networks. *Energy*. **267**, 126460. <https://doi.org/10.1016/j.energy.2022.126460> (2023).
44. Chen, K. et al. A data-driven model of pipe diameter and insulation thickness optimization for district heating systems. *J. Phys. Conf. Ser.* <https://doi.org/10.1088/1742-6596/2166/1/012046> (2022).
45. Çomaklı, K., Yüksel, B. & Çomaklı, Ö. Evaluation of energy and exergy losses in district heating network. *Appl. Therm. Eng.* **24**(7), 1009–1017. <https://doi.org/10.1016/j.applthermaleng.2003.11.014> (2004).
46. Wang, J. et al. Low-order gray-box modeling of heating buildings and the progressive dimension reduction identification of uncertain model parameters. *Energy*. **294**, 130812. <https://doi.org/10.1016/j.energy.2024.130812> (2024).

## Acknowledgements

We would like to acknowledge the language support of Ms. Anna Rohm and especially thank for proofreading. This research was funded by the Stadtwerke Flensburg GmbH and the Society for Energy and Climate Protection Schleswig-Holstein (Gesellschaft für Energie und Klimaschutz Schleswig-Holstein GmbH, EKSH), grant number 8/12-57.

## Author contributions

S.W. carried out the conceptualization, the formal analysis and investigation, developed the methodology and prepared the original draft. M.F. and J.G. contributed to the formal analysis and investigation. D.V. acquired the

www.nature.com/scientificreports/

funding with support of S.W., M.F. and D.V. were in charge of supervision. All authors reviewed the manuscript. All authors have read and agreed to the published version of the manuscript.

### Funding

Open Access funding enabled and organized by Projekt DEAL.

### Declarations

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-15777-0>.

**Correspondence** and requests for materials should be addressed to S.A.W.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025

### 3.2.3 Supplemental Material of Publication 2

Supplementary Information

## Feature Selection for Specific Prediction Targets at the User Level in a District Heating Network

Samanta A. Weber<sup>1,2\*</sup>, Michael Fischlschweiger<sup>1</sup>, Dirk Volta<sup>2</sup> & Jens Geisler<sup>2</sup>

<sup>1</sup> Chair of Technical Thermodynamics and Energy Efficient Materials Treatment,  
Institute for Energy Process Engineering and Fuel Technology, Clausthal University of  
Technology, 38678 Clausthal-Zellerfeld, Germany

<sup>2</sup> Energy and Life Science, University of Applied Sciences Flensburg, 24943 Flensburg,  
Germany

#### Corresponding Author

\*Samanta A. Weber - <https://orcid.org/0000-0002-6487-8248>;

samanta.weber@tu-clausthal.de

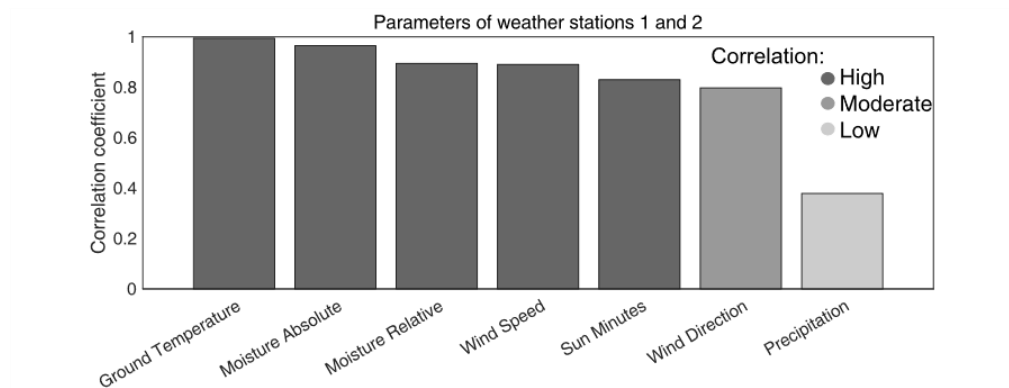
samanta.weber@hs-flensburg.de

**Table S1** Cleaning methods applied to the data collected at the infeed facility

| Parameter at CHP   | Treatment of outliers                                                                                                                                                           |
|--------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Volume flow        | Linear interpolation using mean; previous value using mean; previous value using moving mean with a 30 days window and a threshold factor of 2.5                                |
| Supply temperature | Linear interpolation using mean, threshold factor 3.25; Linear interpolation using moving mean over 1-day window, threshold factor 2.5                                          |
| Return temperature | Linear interpolation using moving median over 71-day window, threshold factor 2.75                                                                                              |
| Accumulated heat   | Linear interpolation using moving mean over 14-day window, threshold factor 0.25; Linear interpolation using moving mean over 16 h-window, threshold factor 0.25                |
| Heat flow          | Linear interpolation using moving median over 10 days-window, threshold factor 2.75                                                                                             |
| Supply pressure    | Linear interpolation using mean                                                                                                                                                 |
| Return pressure    | Linear interpolation using mean                                                                                                                                                 |
| Moving mean        | $MA = \frac{1}{k} \times \sum_{i=n-k+1}^n p_i$ <p>n: last sample in considered window</p> <p>k: window length (number of samples in window)</p>                                 |
| Moving median      | $MM = median(p_i, p_{i+1}, \dots, p_n)$ <p>With <math>i = n - k + 1</math></p> <p>n: last sample in considered window</p> <p>k: window length (number of samples in window)</p> |

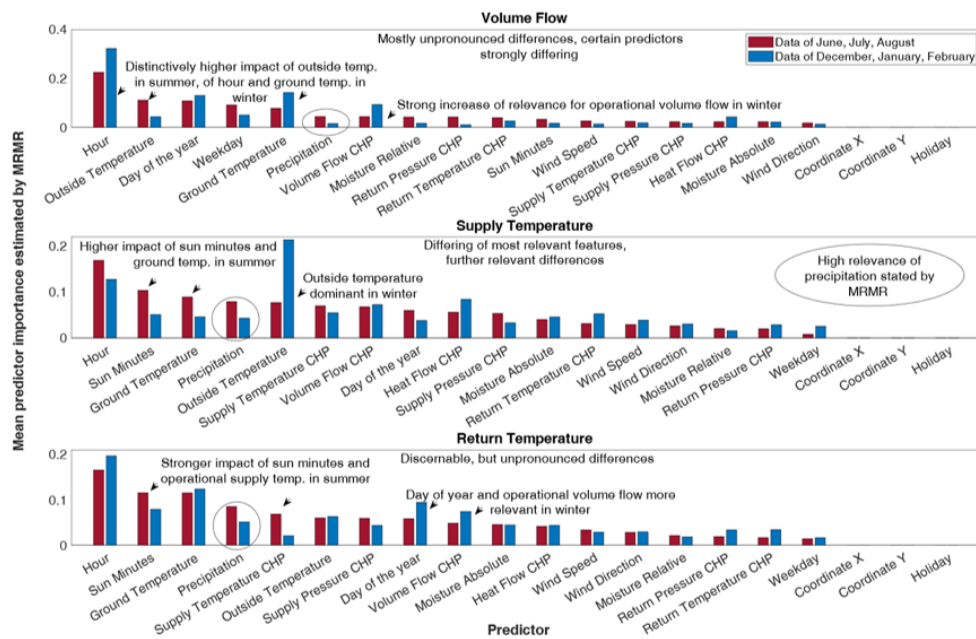
**Table S2** Cleaning methods applied to the meteorological data

| Meteorological parameter | Treatment of                                                                                                           |                                                                                                                                                                            |
|--------------------------|------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                          | Missing values                                                                                                         | Outliers                                                                                                                                                                   |
| Ground temperature       | Linear interpolation with a maximum gap of 72 h                                                                        | Moving median over 500 h and threshold factor of 3.5                                                                                                                       |
| Precipitation            | Linear interpolation with a maximum gap of 3 h                                                                         | Linear interpolation using mean                                                                                                                                            |
| Absolute moisture        | Station 1: Linear interpolation with a maximum gap of 3 h<br>Station 2: Nearest values                                 | Station 1: Linear interpolation using moving median with an 18 h-window, threshold factor 8<br>Station 2: Linear interpolation using 24 days as window for the moving mean |
| Relative moisture        | Station 1: Linear interpolation with a maximum gap of 3 h<br>Station 2: Linear interpolation with a maximum gap of 3 h | Station 1: Linear interpolation using moving median with an 18 h- window, threshold factor 8<br>Station 2: Linear interpolation using mean, threshold factor 4.25          |
| Sun minutes              | Linear interpolation with a maximum gap of 3 h                                                                         | -                                                                                                                                                                          |
| Wind speed               | Linear interpolation with a maximum gap of 3 h                                                                         | Station 1: Linear interpolation using mean, threshold factor 1                                                                                                             |
| Wind direction           | Linear interpolation with a maximum gap of 3 h                                                                         | Station 2: Nearest value using mean, threshold factor of 2.25                                                                                                              |



**Figure S1** Linear correlation coefficients of the parameters of weather stations 1 and 2.

The generally high correlation of the data of both weather stations indicates that the applicability to the third location (model region) is given, apart from the data on precipitation showing a low correlation.



**Figure S2** Results for MRMR, summer, and winter separately. Comparing summer and winter data, the most relevant predictors differ in their evaluation, indicating the necessity to account for the seasons separately. The general order of relevance changes for specific features.

### 3.3 Publication 3: Dynamic User-Level Parameter Modeling in Heat Networks with Deep Learning Long Short-Term Memory

Subobjective 3 is to develop a modeling framework for the suggested modeling targets, which reflect highly dynamic user-level behavior, and to reduce modeling error to an acceptable level by thoroughly optimizing the modeling setup. Existing knowledge of ML-based user-level modeling in heat networks is scarce, particularly for data reflecting the highly dynamic behavior from residential buildings with a limited number of regular occupants, e.g., single-family homes. To achieve subobjective 3, publication 3 integrates the results from publications 1 and 2. It examines the knowledge required to establish an ML-based modeling framework for highly dynamic user-level behavior in the model region's heat network.

First, publication 3 conducts data pre-preparation. This includes cleaning tasks, such as filling missing values, removing outliers, and equalizing timestamps. Methods for scaling and transformation are compared, outlining zero-mean unit-variance scaling and the sine-cosine transformation as beneficial. The comparison is based on the specified error metrics and the achieved runtime. Therefore, this step is intended to enhance modeling performance and reduce computational costs. This type of information contributes to greater modeling success in studies with similar modeling objectives.

Second, publication 3 establishes appropriate hyperparameter ranges for Bayesian optimization. This step helps reduce computational costs and improve model performance by limiting the hyperparameter intervals to be investigated. The results of the hyperparameter optimization step are analyzed to reevaluate the pre-selected hyperparameter ranges. This enables confirming, widening, or narrowing the examined hyperparameter intervals. Therefore, this step helps avoid unsuccessful trials, i.e., hyperparameter combinations that yield insufficient model performance. This knowledge accelerates model development for comparable objectives.

Third, using the pre-preparation methods and the adjusted hyperparameter ranges, domain-specific influences are systematically tested. These include the effects of data quantity and seasonality. Thus, publication 3 provides insights into addressing seasonality and quantifies the required data volume to the equivalent of half a year sampled from all four seasons, thereby helping to define the boundary conditions for modeling user-level parameters in heat networks.

Consequently, publication 3 investigates the scalability of the results within the model region by extending the model to include more substations. Models are built for all substations with a data collection rate of at least 90%. Using the findings from publication 2, the substation data unsuitable for modeling are removed. Lastly, the resulting model is evaluated to highlight the achieved performance. The findings confirm that the modeling error is reduced to a level comparable to benchmarks in the existing literature, despite the high dynamics captured by the modeled user-level parameters.

Overall, publication 3 builds on publications 1 and 2 and presents findings that contribute to a deeper understanding of ML-based modeling of highly dynamic user-level heat network parameters. Information on these user-level parameters, as accessed by modeling and analysis, can improve heat network efficiency and flexibility by enhancing future control setups.

#### 3.3.1 Author Contribution Statement of Publication 3

Weber Samanta: Writing – original draft, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. Dirk Volta: Supervision, Funding acquisition. Michael Fischlschweiger: Supervision, Investigation, Formal analysis.

### 3.3.2 Publication 3

WEBER, S.A./FISCHLSCHWEIGER, M./VOLTA, D.: Dynamic user-level parameter modeling in heat networks with deep learning Long Short-Term Memory. Energy Reports, 14, 2720-2731, <https://doi.org/10.1016/j.egy.2025.09.032>, 2025.

Eq. 2-1 in publication 3 reads as follows in index notation:

$$MeanMaxAbsE = \frac{1}{N} \sum_{n=1}^N \max_j |y_{n,j} - \hat{y}_{n,j}|$$

Publication 3

Eq. 2-1



Contents lists available at ScienceDirect

## Energy Reports

journal homepage: [www.elsevier.com/locate/egyr](http://www.elsevier.com/locate/egyr)

## Research Paper

## Dynamic user-level parameter modeling in heat networks with deep learning Long Short-Term Memory

Samanta A. Weber<sup>a,b,\*</sup>, Michael Fischlschweiger<sup>b</sup>, Dirk Volta<sup>a</sup><sup>a</sup> Energy and Life Science, University of Applied Sciences Flensburg, Flensburg 24943, Germany<sup>b</sup> Chair of Technical Thermodynamics and Energy Efficient Material Treatment, Institute for Energy Process Engineering and Fuel Technology, Clausthal University of Technology, Clausthal-Zellerfeld 38678, Germany

## ARTICLE INFO

## Keywords:

Machine learning  
Automated model building  
Heat decarbonization  
Tailored energy supply

## ABSTRACT

Providing access to user-level information in heat networks is crucial for optimizing control and facilitating the transition to carbon neutrality. Machine learning can automate model development in this area. While models for heat demand have been successfully applied to multi-apartment homes, buildings such as single-family homes pose challenges due to more complex data dynamics. This study aims to develop a Long Short-Term Memory modeling framework that captures high user-level data dynamics, focusing on the parameters of volume flow, supply, and return temperatures, beyond just heat demand. Sensitivity and comparative analysis of modeling errors evaluate data pre-preparation, hyperparameter tuning, and sampling strategies. Initially, data from 15 users are analyzed and rationalized. The model is then scaled up by increasing the number of substations to validate findings. Results show a mean absolute error of 0.0612 for heat demand prediction over 24 h in the upscaled model. Challenges such as the impacts of human behavior at the respective buildings and missing data are discussed. These insights can enhance automation of user-level parameter modeling, improving heat network control and supporting sustainable heat supply.

## 1. Introduction

Heat network modeling aims to improve flexibility and increase the overall efficiency of the energy system (Ntakolia et al., 2022). To achieve this, it is recommended that research should focus more on bottom-up control, integrating knowledge of the heat network at the user level to support reliable heat distribution with adequate temperature safety margins (Mbiydzanyuy et al., 2021). In this context, machine learning (ML) offers the potential to automate the model-building process (Mbiydzanyuy et al., 2021), which could facilitate adaptation to various heat networks. However, in ML modeling, datasets require individual treatment (Leiprecht et al., 2021).

To date, research has strongly focused on predicting the load at the infeed of heat networks (Mbiydzanyuy et al., 2021), for which the required knowledge is provided. For heat load modeling, the literature states explicit information on the ML modeling, for instance, the benefit of resting on the concept of Long Short-Term Memory (LSTM) (Ćirić et al., 2020). In contrast, knowledge of ML modeling for other heat network parameters, particularly at the user level, is limited and needs

further research (Balali et al., 2023). Because of the potential to improve heat network control, user-level parameter modeling is currently being explored, though it is still in the early stages (e.g., Idowu et al., 2014a). Challenges in predicting heat demand for highly dynamic user-level data are highlighted (Idowu et al., 2016). As one of the existing studies, Idowu et al. (2014b) reports a mean absolute percentage error (MAPE) of  $21.5455 \pm 5.71\%$  for the prediction target of heat demand and a 24-h-ahead prediction for a building with twelve apartments using Support Vector Regression. It is noted that the modeling error increases when modeling the heat demand of buildings with fewer regular occupants (Idowu et al., 2014b). Idowu et al. (2014b) explores alternative ML architectures, including regression trees, feedforward neural networks, and multiple linear regression.

However, a specific knowledge gap exists in combining user-level parameters of heat networks as modeling targets, especially for the highly dynamic user level of residential buildings with few occupants, and the promising modeling architecture of the LSTM so far.

Hence, this study focuses on the development of LSTM-based frameworks for modeling user-level parameters in heat networks, fully considering the dynamic behavior of individual residential buildings

\* Corresponding author at: Energy and Life Science, University of Applied Sciences Flensburg, Flensburg 24943, Germany.

E-mail addresses: [samanta.weber@tu-clausthal.de](mailto:samanta.weber@tu-clausthal.de), [samanta.weber@hs-flensburg.de](mailto:samanta.weber@hs-flensburg.de) (S.A. Weber).

<sup>1</sup> <https://orcid.org/0000-0002-6487-8248>.

<https://doi.org/10.1016/j.egyr.2025.09.032>

Received 5 May 2025; Received in revised form 14 August 2025; Accepted 20 September 2025

Available online 27 September 2025

2352-4847/© 2025 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

### Nomenclature

|             |                                              |
|-------------|----------------------------------------------|
| AE          | Absolute error                               |
| MAE         | Mean absolute error                          |
| MAPE        | Mean absolute percentage error               |
| MeanMaxAbsE | Mean maximum absolute error                  |
| MSE         | Mean squared error                           |
| N           | Number of samples                            |
| RMSE        | Root mean squared error                      |
| RSE         | Square root of squared error                 |
| SE          | Squared error                                |
| y           | Measured (test) value of the target          |
| $\hat{y}$   | Model evaluation for the value of the target |

such as single-family homes with few regular occupants. The further goal is to reduce the error of such automatically developed machine learning models to an acceptable level. It is hypothesized that thorough optimization of manual modeling steps—through comprehensive testing and analysis of model settings—can achieve error margins comparable to those reported in existing literature, although these studies involve buildings with more regular occupants and thus less dynamic behavior. The manual steps examined include data scaling and transformation, hyperparameter tuning, and careful selection of training, validation, and testing datasets, considering seasonality and data size. Unlike typical heat demand modeling, this work targets the volume flow, supply temperature, and return temperature at the building level. These parameters are the essential ones for calculating heat flow, which depends on volume flow, specific heat capacity, mass density, and temperature differences between supply and return. Furthermore, the selected targets also represent supply safety margins and heat efficiency, making them valuable for improving heat network management.

The study's methodology focuses on optimally combining highly dynamic user-level parameters with an LSTM network architecture. Feature selection for the modeling targets was addressed in a prior study (Weber et al., 2025), and those findings are considered in this work. The process begins with data from 15 substations that have suitable characteristics, providing essential information for data pre-preparation, automated hyperparameter tuning, and domain-specific selection of data samples. Sensitivity and comparative analyses of the error metrics of the automatically developed models identify the most effective choices for the investigated settings. To validate these findings, models are subsequently generated for substations with at least a 90 % data collection rate, demonstrating the scalability of the information within the heat network. The heat demand is also calculated and evaluated as a modeling target to compare with existing research. A 24-h-ahead prediction model is then developed for benchmarking, highlighting ongoing challenges and opportunities within the study's scope. The success of these settings could inform related research, helping to accelerate and improve the automation of ML model development within this respective field.

The work is organized as follows: Section 2 describes the research methodology and model development. Section 3 presents and analyzes the results, starting with sensitivity and comparative analyses for the 15 selected substations for rationalization purposes and model testing. This is followed by validating the insights gained by training and evaluating nowcast and 24-h prediction horizon models for all substations with data collection rates of at least 90 %. The most successful models are compared with the literature using the overall and instantaneous error metrics. Section 4 outlines the conclusions and implications of the findings for potential directions of future research.

## 2. Material and methods

### 2.1. Framework concept, model region, data acquisition and cleaning

A conceptual framework for data processing was developed to support the aim of this study, as shown in Fig. 1a). The data are collected in the model region depicted in Fig. 1, which illustrates the general piping structure of the heat network (b), blue lines) and the locations of substations with data collection rates of 90 % or higher (c) dark blue dots).

The processing steps in this study are highlighted in dark blue in Fig. 1a). The main idea is to forecast inputs and provide nowcasts with the model. In Fig. 1a), recently recorded data are shown as a solid line, while historical data are depicted with a broken line. The ML models use historical data as inputs for training. This data originate from the centrally supplied district heating system in Tarp, northern Germany, which has 486 substations (Fig. 1b, c). In 2023, the heat input was 26 GWh, with a peak load of 12 MW. The minimum design supply temperature for user substations is 70 °C when outside temperatures are  $\geq 15$  °C, increasing linearly up to 100 °C when outside temperatures are  $\leq -10$  °C (Technische Anschlussbedingungen (TAB), 2011). The minimum infeed supply temperature is 70 °C, and the maximum is 110 °C for temperatures  $\leq -10$  °C (Technische Anschlussbedingungen (TAB), 2011). The connected substations mainly serve single-family homes or apartment buildings with few units. Public buildings, such as schools, gyms, a church, and a large residential complex connected to the same substation (top left corner of Fig. 1b), and one industrial user are also connected. The substation data were recorded hourly from December 1, 2021, to January 15, 2024. The data should be clean and have no outliers. Additionally, the selected data should neither exhibit a fault that impacts the recordings nor a significant change in behavior. Preferably, the substations should share distribution pipes to enable the LSTM algorithm to capture the commonalities. Only the substations marked in Fig. 1c) are included in the study because data collection rates below 90 % are unsuitable for modeling. All selected substations will be modeled in parallel since creating separate models for each substation becomes less practical as the number of substations increases.

The inputs for these models include operational data from the heat network infeed facility, meteorological data, and temporal data. Table 1 summarizes the datasets used for feature engineering, as described in Weber et al. (2025).

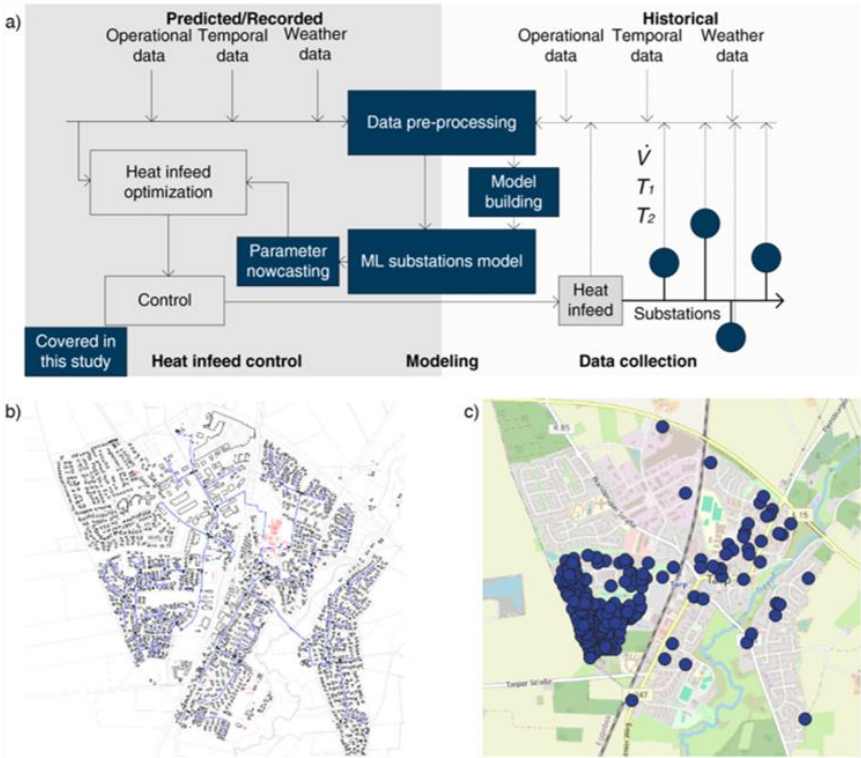
The operational heat flow is designed as a feature of the accumulated heat by calculating the differences between consecutive measurements. Although the operational data and outside air temperature recordings originate from the heat infeed facility, this study uses measurements from the German Meteorological Service (Deutscher Wetterdienst) (Index of climate.environment/CDC/observations\_germany/climate/hourly/, 2023) at the Leck weather station for the remaining meteorological parameters, as no weather data are collected directly at the model region's location. The Leck weather station is the nearest one aligned with the most common wind direction. Unlike Weber et al. (2025), precipitation is excluded from the feature set due to the potential uncertainty associated with this meteorological data, which results from the change in location.

### 2.2. Model training, hyperparameter optimization, and model assessment

#### 2.2.1. Long Short-Term Memory architecture and hyperparameter optimization

The existing literature on the particular field of data-driven heat load modeling currently names Deep Neural Networks in general and LSTM specifically as highly sophisticated (Čirić et al., 2020). Regarding the so-far successful application of the LSTM in heat load prediction, optimizing the combination with the user-level data can significantly enhance the modeling performance.

Thus, this study employs LSTM frameworks. For solving sequential data problems like time series, recurrent structures work by storing



**Fig. 1.** General framework of this study and network structure of the model region. a) gives an overview of the conceptualization modeling framework of this study, b) depicts the piping structure thermal energy system (Weber et al., 2024) (blue lines), and c) the location of substations with data collection rates  $\geq 90\%$  (dark blue dots).

**Table 1**  
Datasets for feature engineering from Weber et al. (2025).

| Internal features  | External features   |                         |
|--------------------|---------------------|-------------------------|
| Operational        | Meteorological      | Temporal (behavioral)   |
| Supply temperature | Outside temperature | Hour of the day         |
| Return temperature | Relative moisture   | Day of the week         |
| Volume flow        | Sun minutes         | Day of the year         |
| Heat flow          | Wind speed          | Occurrence of a holiday |
|                    | Wind direction      |                         |

information from previous steps. This enables the model to reflect long-term dependencies. As a type of recurrent structure, the LSTM uses a loop that returns processed information from the sequential inputs (Kemper et al., 2025).

Fig. 2 depicts the LSTM layer structure of this study schematically. Apart from the LSTM layers, a minimum of additional layers is required, as illustrated in Fig. 2.

As shown in Fig. 2, the sequence input layer receives the structured features and passes them along. Stacked LSTM layers follow, with the

number of layers determined during hyperparameter optimization. The number of parallel cells within these layers is also optimized during this process. The first fully connected layer takes the output of the LSTM layers and passes on the weighted results. The subsequent dropout layer regularizes by randomly setting nodes received from the first fully connected layer to zero at a specified rate (dropout rate) to reduce overfitting. It then passes the remaining information to the second fully connected layer. The number of neurons here matches the number of outputs (targets), allowing the regression layer to provide estimates for the modeling targets. The training loss function is the mean squared error (MSE), as used effectively in related literature (Houran et al., 2023).

This study details the training setup for an LSTM model, emphasizing reproducibility by providing specific configurations in Table 2 (Supplementary material). Default parameter settings are applied where not explicitly mentioned, following existing examples (MathWorks, 2024). The importance of data shuffling for accounting for seasonal effects must be highlighted. Specifically, this work performs a single shuffle before data splits. Additionally, this work sorts sequences by length prior to batch formation to minimize padding.

Next, the model hyperparameters are selected for the architectural



**Fig. 2.** Schematic architecture chart for the LSTM applied in this work. The figure captures the sequential order and type of the layers.

setup. If an ML architecture is applied with sub-optimal hyperparameters, the result will not be the algorithm's achievable optimum (Song et al., 2021). Standard methods of automated hyperparameter tuning include grid search, random search, and Bayesian optimization (Jiang et al., 2022). Bayesian optimization was found to identify more sophisticated hyperparameters than a human expert and to be faster than alternative methods (Snoek et al., 2012). It begins with a random selection within pre-defined intervals (Snoek et al., 2012) and searches the parameter space to find the global minimum (Kim et al., 2018). Kim et al. (2018) significantly reduces the number of required training cycles with Bayesian optimization by utilizing parameter intervals determined as suitable for similar datasets. Therefore, Bayesian optimization is used for hyperparameter tuning specific to the modeling targets in this study. Table 2 (Supplementary material) details the hyperparameters being tuned, including their intervals, input types, and transformation methods.

The Bayesian tuning process involves training a total of 1000 networks, which represent the iterations in the optimization. The primary goal is to minimize the MSE, consistent with related studies such as Song et al. (2021). Unlike the training phase, hyperparameter optimization evaluates the MSE on test data, providing both the trained networks and their respective MSE values for test and validation datasets. This approach facilitates selecting a network that balances performance across both datasets. To optimize computational efficiency and improve the chances of finding the best hyperparameter settings, the initial parameter intervals are re-evaluated using histograms in Section 3.1.2. These histograms indicate the success of the interval choices, allowing adjustments to the hyperparameter ranges as necessary.

### 2.2.2. Error metrics and evaluation strategy

The modeling performance is assessed using separate data not applied in the training or hyperparameter optimization process, to avoid information leakage from test data to training results. To evaluate the results, this study employs standard error metrics: the MSE, the root mean squared error (RMSE), and the mean absolute error (MAE). Furthermore, this work utilizes the mean maximum absolute error (MeanMaxAbsE), as explained below: As is commonly known, one key characteristic of an ML model is a sophisticated balance between variance and bias. The MeanMaxAbsE can aid in estimating how well the model performs in this regard, as given in (2-1) (MathWorks, 2023).

$$\text{MeanMaxAbsE} = \frac{1}{N} \bullet \sum_{n=1}^N \max \left( + \sqrt{(y_n - \hat{y}_n)^2} \right) \quad (2-1)$$

The MeanMaxAbsE measures the modeling target with the most significant deviation between estimation and test data for each instance. This specific target is typically the most difficult to estimate, potentially due to high data variability. Therefore, the MeanMaxAbsE tends to be considerably higher when the model isn't flexible enough to capture the data dynamics fully. As a result, a high MeanMaxAbsE is an indication that a model with a better balance between variance and bias should be preferred.

The error metrics illustrate how the selected settings influence the model's performance. Various modeling architectures will be generated through Bayesian hyperparameter optimization, allowing comparison based on error margins using the specified metrics. This process quantifies the effect of particular model configurations on the results, following the principles of sensitivity analysis. It provides clear insights into the success of different settings within the modeling process. The results of this comparison are detailed and discussed in Sections 3.1 and 3.2.

In addition to overall performance metrics, this study evaluates the model's seasonal dependence using instantaneous error measures: squared error (SE), maximum absolute error (MaxAbsE), root mean squared error (RMSE), and absolute error (AE), as discussed in Section 3.2.

### 2.3. Data from 15 substations for model settings' assessment

15 substations were identified as the most suitable for providing measurements for a first assessment of the model-building process. Fig. 3 (Supplementary material) illustrates the min-max scaled data for volume flow, supply temperature, and return temperature across 15 selected substations during a typical winter and summer week. This visualization provides insights into data patterns and dynamics. Notably, the volume flow is lower and demand occurs less frequently in summer compared to winter. Additionally, temperature levels vary between seasons, with the supply temperature readings in summer exhibiting greater apparent randomness than in winter.

#### 2.3.1. Data pre-preparation

After acquiring the data, the raw data must be pre-prepared. First, for the substation data, missing values are filled using a moving mean, and sample intervals are standardized through linear interpolation. Due to the high volatility of the data, a maximum window of three hours is applied when filling gaps. Next, for the machine learning inputs, Table 3 (Supplementary material) outlines methods for outlier removal in the operational dataset. Unlike many studies, Chujai et al. (2013) recommend using the previous value to fill in missing data, a method also adopted here. The data are then resampled to hourly intervals via linear interpolation. For meteorological data, wind speed outliers are corrected by linear interpolation based on the mean, timestamps are aligned by interpolation, and missing measurements are filled with the previous value. Finally, temporal features are extracted from timestamps, including a binary indicator for holidays, where 1 indicates a holiday and 0 indicates a regular day (Potočník et al., 2021).

The next step in data pre-preparation in this study is scaling. This aims to transform the numeric values of the features into a generally valid range to prevent the influence of magnitudes (Singh and Birmohan Singh, 2020). A common step is first shifting the mean value to zero (e.g., Hou et al. (2022)). Therefore, this is applied to all non-binary, non-cyclic features. Next, this study tests scaling with different statistically assessed properties. Existing scaling methods include standardization, min-max normalization (Singh and Birmohan Singh, 2020), and unit variance (Greff et al., 2017). This work refers to standardization as scaling by standard deviation and unit variance as scaling by variance. Min-max normalization uses the range between the maximum and minimum values. Regarding the encoding of cyclic continuous data, for a linearly increasing value, a step in the time series occurs at the end of each completed cycle. Thus, the literature highlights the superiority of transforming specific cyclic continuous features (such as features for seasonality) with sine and/or cosine (Potočník et al., 2021). Sine-cosine transformation generates equal distances between all consecutive time steps (Mahajan et al., 2021). The transformation divides each temporal value (e.g., hour of the day) by its range (e.g., 24 h) and applies the sine or cosine to this quotient, multiplied by  $2\pi$ . Section 3.1.1 conducts a sensitivity analysis of the choice of scaling and transformation methods by comparing the resulting values of the MSE and MeanMaxAbsE, as well as the runtime for training the models. Furthermore, the flexibility of the model is investigated based on model evaluations compared to the test data.

#### 2.3.2. Data quantity and seasonality

With the combination of modeling architecture and data pre-preparation completed, domain-specific influences should be compared. For the modeling targets of this study, the impacts of data quantity and seasonality will be investigated. Generally, data size has a significant effect on the success of modeling. In the context of heat network modeling, Johansson et al. (2017) demonstrate that a more extensive training set does not always directly correlate with higher modeling performance, as it carries the risk of increased overfitting. Furthermore, the literature emphasizes the impact of data seasonality on heat networks (Laakkonen et al., 2017). This is because the demand for

space heating, which dominates during the heating period, can be predicted more precisely than the demand for domestic hot water, which often forms the sole type of demand in summer and is strongly influenced by human behavior (Bujalski et al., 2023).

Therefore, models are optimized for different training and testing datasets. In all cases, the test data comprise 10 % of the total number of samples. The first set, originating from winter, is formed from 5 % of the total data, taken from the beginning and end of the measurement period. The second set centers around mid-summer (July 15, 2022). The third set includes winter 2022 (January 15), spring (April 15), summer (July 15), and fall (October 15), with each season contributing a quarter of the overall test data. Subsequently, three different sizes of training data are used to identify the optimal amount of data for this study. The first set contains 15,799 samples (excluding test data), the second includes 8760 samples (one year), taken from the most recent data after the test set was set aside, and the third comprises 4380 samples (half a year). For this number of training samples, data from the first and second semesters, as well as from the middle of each season, are used. The MSE and MeanMaxAbsE serve as the basis for assessing model performance in Section 3.1.3, through a sensitivity analysis of how input data seasonality and quantity affect the modeling error.

#### 2.4. Data of substations with at least a 90 % data collection rate for validation

The assessment results for 15 substation models will be validated by constructing models of additional substations within the heat network. These substations must have a data collection rate of at least 90 % (see Fig. 1c)).

##### 2.4.1. Settings for validation models

The following addresses the necessary adjustments for modeling a larger number of substations. Specifically, the concept for treating missing values for gaps exceeding three hours must be altered, as removing across all inputs and outputs would diminish the dataset's representativeness of network behavior. As the input dataset is mainly complete, all instances with missing values are deleted. However, missing samples in the modeling targets are filled using a different strategy. The study tested three approaches: filling with the previous value (*Technische Anschlussbedingungen (TAB)*, 2011) (resulting in an MSE of 0.0430 on evaluation data), filling with zero in unscaled data (MSE: 0.0524), and filling with  $-1$  in scaled data (MSE: 0.0423). The lowest MSE was achieved by replacing with  $-1$ . Therefore, for the extended model, missing values in the scaled target data are filled with  $-1$ .

Apart from filling in missing values, it is essential to assess whether the collected datasets accurately represent the overall behavior of substations. Notably, for certain substations, the recorded patterns show significant changes during the measurement period, often due to faults or anomalies. An example of such a dataset is depicted in Fig. 4 (Supplementary material). These disruptions pose challenges, mainly if they occur only in the evaluation data, because the model cannot learn these patterns during training. To address this, a three-step process is proposed: First, train an optimized model for substations with at least 90 % data collection completeness (192 substations). Second, exclude substations with detected discontinuities from the dataset for a second evaluation, leaving 166 substations. Third, develop a separate algorithm tailored for these 166 substations. The performance of these models was compared using global error metrics, as detailed in Section 3.2.1.

##### 2.4.2. Validation models' assessments

This section describes the visualization methods used to evaluate the modeling performance. In Section 3.2.2, a comparison is made between the nowcasts and test data, focusing on the mean values across all modeling targets and substations within the evaluation dataset. The most sophisticated model is highlighted to give an initial impression of

the modeling success. To further assess performance, the current values of selected error metrics will also be visualized. This approach allows for the identification of seasonal patterns in the errors, such as periods with increased or decreased errors, and understanding how error variance changes over the year.

For the heat demand modeling discussed in Section 3.2.3, the three modeling targets of the volume flow, the supply, and the return temperatures are combined. A visualization provides an initial overview, illustrating the dynamics of the error metrics over the modeled period. These analyses aim to determine whether the individual errors of the modeling targets tend to stabilize or intensify, and to assess the seasonal influence.

The comparison between the current study results with the LSTM-based framework combined with highly dynamic user-level parameters and existing literature benchmarks regarding heat demand prediction with lower user dynamics is discussed. The MAPE of up to  $21.5455\% \pm 5.71\%$  reported in Idowu et al. (2014b), for a 24-h-ahead heat demand forecast for a Swedish building with a low apartment count (twelve units) uses a model trained on data from January to April 2014. In this study, the MAPE cannot be calculated for the modeling targets due to zero readings, which lead to infinite values during calculation. To facilitate comparison, the study's output data are normalized using min-max scaling to a range of 0–1, aligning the RMSE and MAE results with percentage errors. A model with a similar setup to the nowcasts is trained and evaluated for a 24-h forecast horizon, as detailed in Section 3.2.3, allowing for direct comparison with the nowcast predictions and test data. Additionally, one-day sequences during summer and winter are used to illustrate how an increased prediction horizon impacts modeling performance.

### 3. Results and discussion

#### 3.1. Model settings' assessment

##### 3.1.1. Investigation of input scaling and transformation methods

The models for the 15 substations trained through hyperparameter optimization using Bayes' principle are evaluated in Fig. 5. This figure shows the MSE in dark blue and the MeanMaxAbsE in grey-blue for the models that achieved the lowest testing MSE. It also presents different combinations of input data scaling and cyclic continuous feature transformation methods. Specifically, Fig. 5a) illustrates the sensitivity analysis for models utilizing the sine-cosine transformation, while Fig. 5b) displays results for models applying a sawtooth shape to the cyclic continuous features.

The lowest overall MSE (Fig. 5a), dark blue) was achieved with standardization, while the lowest MeanMaxAbsE (Fig. 5a), grey-blue) was observed with unit variance scaling. As shown in Fig. 5, the metric values are generally higher for the sawtooth-shaped temporal inputs. An exception is the MeanMaxAbsE (Fig. 5b), grey-blue) for the min-max normalization.

The computational times varied considerably across different configurations. On a device with 24 cores and 32 processors, training for 1000 cycles took 67 h when using min-max normalization combined with sawtooth-shaped features, whereas it took 92 h with sine-cosine transformed inputs. On a second device with 10 cores and 20 processors, each training cycle averaged 30 min with unit variance or standardization and sawtooth-shaped inputs, compared to only 4 min with sine-cosine inputs. Consequently, sine-cosine transformation is favored for both model performance and computational efficiency, and all subsequent models were trained using this method. The results demonstrate that sine-cosine transformation enhances the performance of models utilizing cyclic continuous features within an LSTM architecture.

To assess the models' variance and bias, Fig. 6 (Supplementary material) provides an overview of the nowcasts from the optimized models for all substations where the input data were scaled with either

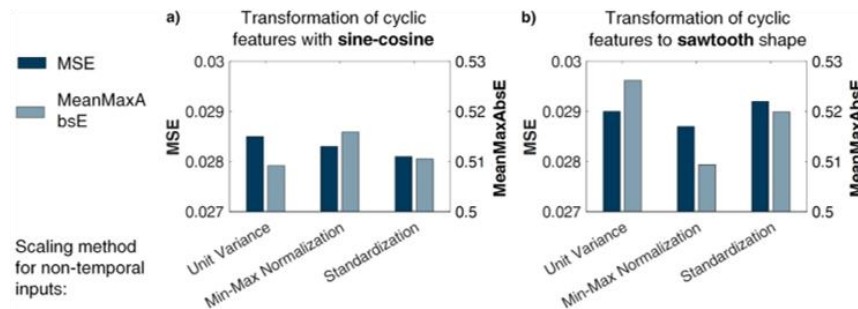


Fig. 5. Sensitivity analysis results for MSE (dark blue) and MeanMaxAbsE (grey-blue). The metrics are accessed with Bayesian hyperparameter optimization for the input data transformation of the cyclic continuous features with a) sine-cosine and b) sawtooth shape.

unit variance (dark blue line), standardization (red line), or min-max normalization (orange line). These are compared to the test data (grey line). To improve clarity, Fig. 7 compares two selected substations, chosen for their contrasting characteristics. Substation No. 2 (Fig. 7a)) exhibits a typical pattern indicative of a strong influence from seemingly random human behavior, along with significant periods of stagnation for

primary supply parameters, which a thermal storage facility enables. Conversely, substation No. 8 (Fig. 7b)) shows reliably repeating patterns, suggesting periodic usage and potentially low influence of human behavior on the supply strategy. Fig. 7 also includes magnified views of specific intervals.

The substation in Fig. 7a) frequently shows zero volume flow

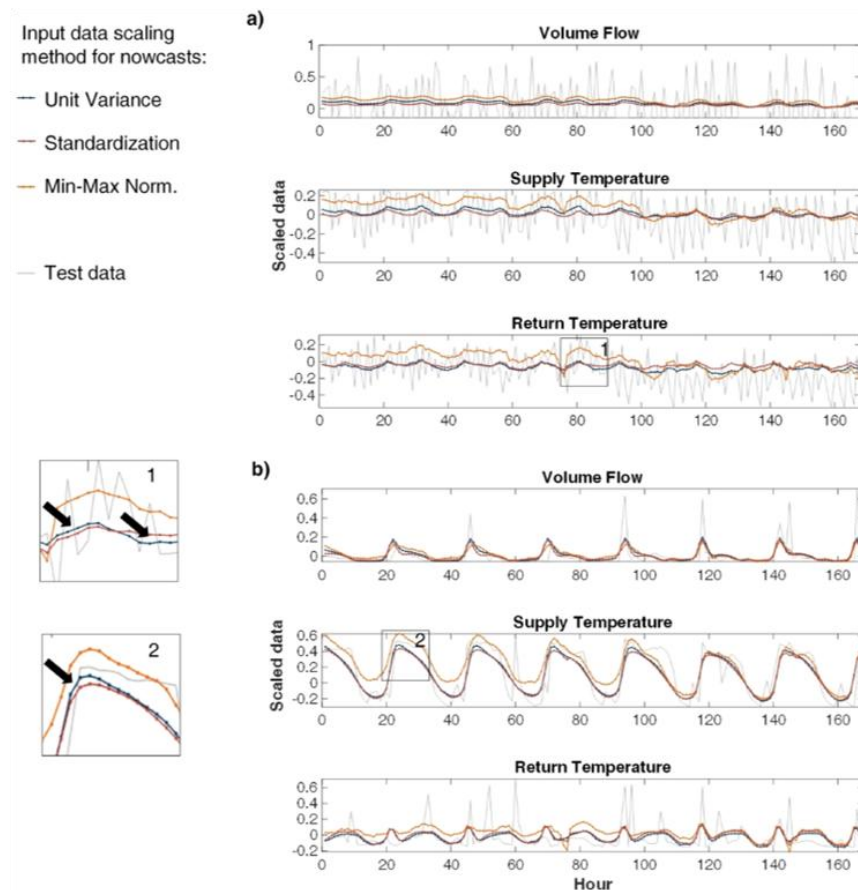


Fig. 7. Nowcasts of optimized models for two exemplary substations. The models were trained applying unit variance (dark blue line with round markers), standardization (red line with round markers), and min-max normalization (orange line with round markers) for scaling and are compared to the test data (grey line) for substation 2 (a)) and substation 8 (b)).

periods. The data contain a high level of randomness caused by human decisions interacting with storage capacities. Thus, the models in Fig. 7a) only reliably capture these dynamics to a limited extent. Conversely, all models in Fig. 7b) effectively represent the dynamics. However, Fig. 7b) suggests an underestimation of both volume flow and return temperature during high-demand periods.

Min-max normalization performs the worst, leading to overestimation during weekdays (Fig. 7, orange line with round markers). Fig. 6 (Supplementary material) indicates a significant bias of min-max normalization for most substations. The model trained with unit variance scaling (dark blue line with round markers) better captures the dynamics than standardization (red line with round markers), as shown in Fig. 7, magnifications 1 and 2. This improved performance results in a lower MeanMaxAbsE (Fig. 5a), grey-blue) for this scaling method.

Overall, the best balance between model bias and variance appears with unit variance scaling. This is confirmed by the sensitivity analysis in Fig. 5 and the comparative visualizations in Fig. 6 (Supplementary material) and Fig. 7. Both the runtime and the sensitivity analysis of error metrics on test data point to superior results with sine-cosine transformation of cyclic continuous features. Since this combination produces the lowest modeling errors relative to this study's goal, all models hereafter are trained using these pre-preparation methods. Future research could explore how well these findings generalize regarding data dynamics and their applicability to other heat networks.

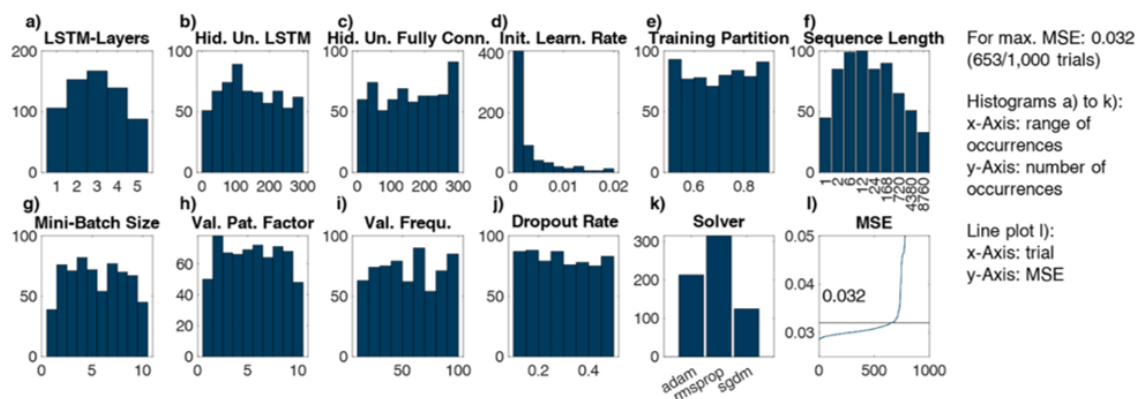
### 3.1.2. Validation of chosen intervals for hyperparameter optimization

The trained models were examined to validate the selection of intervals for hyperparameter optimization using Bayes' principle. To evaluate the hyperparameter choices, histograms (Fig. 8a)–k) serve as a standard method incorporated into this study's modeling framework. Comparing the initially selected intervals based on the probability of choosing a specific setting highlights the importance of the selected intervals, which is further explained below in relation to the results obtained. As this study aims to define intervals for targeted model optimization, trials with an excessively high MSE should be excluded from analysis. Clearly, the combination of hyperparameters used in these trials was ineffective. To identify an appropriate threshold for acceptable MSE, Fig. 8l) displays the MSE values of 1000 models in ascending order. This approach reveals a notable increase above an MSE of 0.032 (Fig. 8l)). This suggests that models with  $MSE \leq 0.032$  perform significantly better than the others. Consequently, models with an MSE

above 0.032 are considered to have been generated with unsuccessful hyperparameter choices. Therefore, only the 653 models with an  $MSE \leq 0.032$  are included in the histograms shown in Fig. 8a) to k).

Generally, the results in MSE for most trials indicate that selecting intervals for hyperparameter initialization based on related studies was successful. Starting with the number of LSTM layers and hidden units, the shapes of the distributions in Fig. 8a) and Fig. 8b), with slopes on both sides, suggest that the initial intervals were appropriate. In contrast, the results for the fully connected layer in Fig. 8c) show successful training at the highest values, which may depend on inference with the dropout rate, so this interval should not be changed. Moving on to the initial learning rate in Fig. 8d), the right-skewed distribution aligns with the expected value for a logarithmically scaled parameter. As few successful trials occur above 0.015, the upper limit of this parameter should be reduced to that value. For the training data partition in Fig. 8e), the distribution appears nearly uniform. As this indicates a limited impact of this hyperparameter, the range is expanded toward higher values, specifically 0.95, to allow the setup to utilize more data for training and potentially improve performance. Concerning sequence lengths in Fig. 8f), the maximum should be lowered to 720 h. The minimum could be increased to two, but since the trial with the lowest MSE used a sequence length of one, the interval remains unchanged. For the mini-batch size and validation patience factor in Fig. 8g) and Fig. 8h), the slopes on both sides suggest that the chosen ranges are suitable and should remain the same. Regarding the validation frequency in Fig. 8i), the skewness is only evident to the left. Extending the range to higher values risks missing the optimal iteration for early stopping and may lead to overfitting, which should be avoided. For the dropout rate in Fig. 8j), values above 50 % likely do not improve the model's performance, even though the distribution shows no clear slope on either side. This interval is also not adjusted, considering the relationship between dropout rate and the number of neurons in the fully connected layer. Finally, in Fig. 8k), although the solvers vary in the number of trials used, all three appear to lead to convergence during training. Therefore, all three solvers will remain in the setup.

While the results show that some models were built with unsuccessful hyperparameter choices, overall, Bayesian hyperparameter optimization produces convincing results for the other models. Additional analysis could help narrow these intervals to speed up model development and reduce the chances of unsuccessful hyperparameter combinations, aligning with the study's goal to minimize modeling



**Fig. 8.** Comparison of histograms for hyperparameter selections of models with  $MSE \leq 0.032$ . The histograms for the hyperparameter settings can assist in determining whether the intervals assessed in Bayesian optimization were chosen adequately for a) the number of LSTM layers, the choice of number of hidden units for b) the LSTM layers and c) the fully connected layer, d) the initial learning rate, e) the portioning of training and validation data, f) the selected maximum possible sequence length, g) target number for splitting into batches, h) the validation patience factor for early stopping of the training and i) the associated validation frequency, f) the dropout rate for the corresponding regularization layer, k) the choice of solver, where an  $MSE \leq 0.032$  implies a successful model setup, as determined with l) the ordered values of the achieved MSE values.

errors. Future research might explore how well these parameter ranges apply to other heat networks and different modeling architectures.

### 3.1.3. The impact of the training data quantity and seasonality on the model performance

The influence of training data quantity and seasonality on model performance is analyzed. Fig. 9 displays the sensitivity analysis for MSE (dark blue) and MeanMaxAbsE (grey-blue) for the test set of winter data (Fig. 9a)), summer data (Fig. 9b)), and the four seasons (Fig. 9c)).

The overall analysis indicates that modeling the winter season results in the lowest error metric values. Specifically, Fig. 9a) shows that the MSE (dark blue) and MeanMaxAbsE (grey-blue) increase when less training data are used. Training on data from the second half of the year yields lower deviations than using data from the first half, with mid-season sampling performing comparably to full-year data regarding MSE. In Fig. 9b), similar trends are observed: training with the second semester results in lower MSE and MeanMaxAbsE than the first semester. Interestingly, the MeanMaxAbsE for all data exceeds that of only the second half, although the MSE remains lower. Testing on summer data shows metrics that fall between the half-year sampling methods, but using 8760 samples results in higher metric values than just the second semester, as seen in Fig. 9b). The results in Fig. 9c) mirror this pattern, indicating that increased data quantity may lead to overfitting, which could be suboptimal, especially for summer and four-season test datasets. To effectively capture the year's general behavior, training data should encompass all seasons, a conclusion supported by this sensitivity analysis. The findings suggest that a minimum of half a year's data is necessary to reduce modeling error adequately. Future research might explore whether incorporating more season-specific data or employing alternative strategies like online learning could further improve model performance.

### 3.2. Validation models' assessment

#### 3.2.1. Models' building and comparison of global error metrics

To validate the acquired knowledge, the following presents the assessment of models for all substations with a data collection rate of at least 90 %. The models were trained and tested using data from December 1, 2021, to December 31, 2022. The training data, sampled around mid-season, covers approximately six months. Data from 2023 were used solely for model evaluation.

For the three extended model configurations—which handle anomalies in substation data differently—described in Section 2.4.1,

Fig. 10 (Supplementary material) presents the calculated metrics and their standard deviations. These metrics are shown for the average of all modeling targets, individual targets, and, for comparison, the min-max scaled heat flow derived from three parameters. Missing samples in the test data were excluded from metric calculations. The results indicate that heat demand at the substation level is generally the most accurately modeled parameter, followed by supply temperature.

Table 4 (Supplementary material) shows the calculated metrics. The results indicate that excluding substations with anomalies in the measured pattern improves modeling success, as reflected in all the metrics used. Comparing the options of excluding these substations before versus after training, the model trained on 192 substations and then removing anomalous patterns achieves better performance across all metrics, except for a small margin in MeanMaxAbsE.

The study's objective was to evaluate the performance in modeling heat flow, as indicated by an MAE of 0.0600 and an RMSE of 0.1071 (see Table 4 (Supplementary material)). The results were obtained by excluding substations with anomalous patterns. Despite high data variability, the model's performance is comparable to the reference MAPE value of 21.5455 % reported by Idowu et al. (2014b) for heat load prediction over a 24-h horizon. This demonstrates the model's robustness, even with highly dynamic data, indicating the potential for further improvements. Based on these findings, selecting an LSTM architecture combined with user-level data appears effective. Extending the model to include substations with at least 90 % data coverage could demonstrate its scalability within the same model region.

#### 3.2.2. Evaluation of modeling targets

This study first examines the overall performance of the mean across all targets. To understand the dynamics and potential seasonality of the evaluation metrics, the performance at each moment—focusing on the setup that yields the best overall results after training on 192 substations and excluding anomalous data—is analyzed. Fig. 11 a) shows the nowcast as the average across all targets and substations compared to observed test data. Additionally, Fig. 11 b)–e) display plots of the SE, MaxAbsE, RSE, and AE for the instantaneous mean across all substations and targets.

In Fig. 11 a), the typical curve of heat demand is visible. Generally, the variance in the model estimations appears higher during the summer period. Thus, Fig. 11 b)–e) indicate an increase in errors during the summer, with a peak between the hours of 5000 and 6000. In comparison, the heating period seems to be modeled more accurately, resulting in generally lower error values in Fig. 11 b)–e). However, when

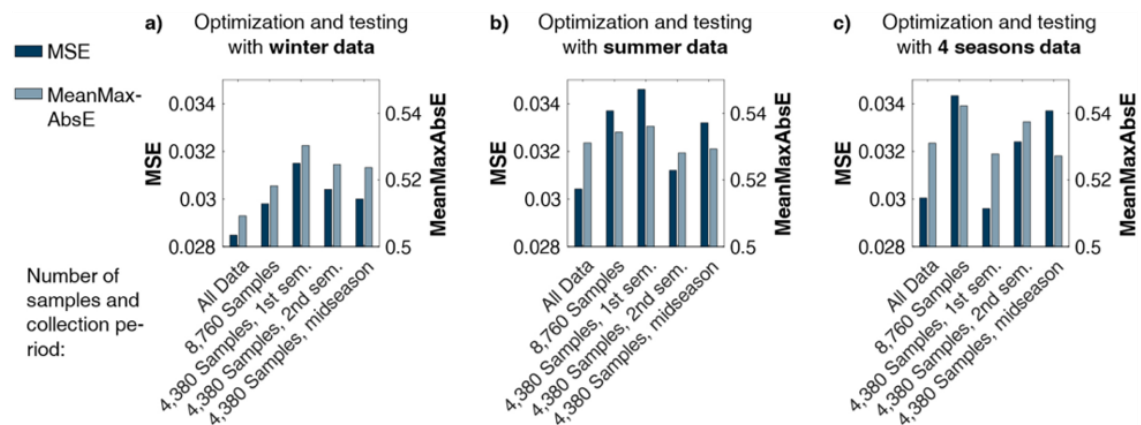
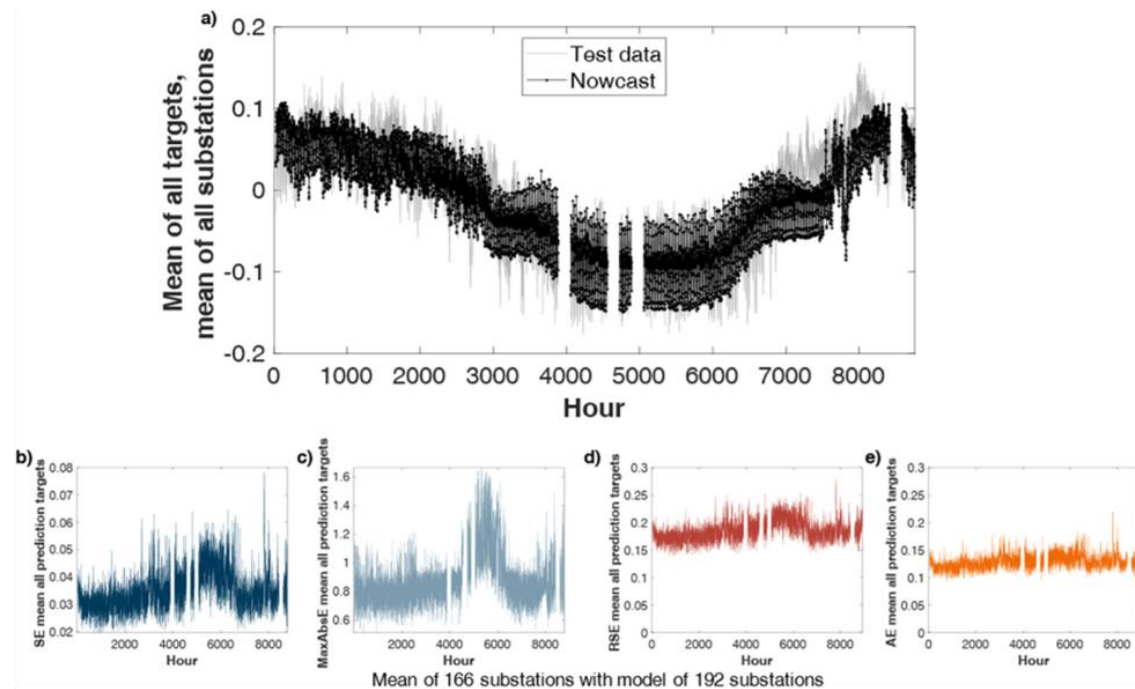


Fig. 9. Sensitivity analysis of model metrics for test data originating from different seasons and different amounts of training data. The figure captures the MSE (dark blue) and MeanMaxAbsE (grey-blue) of the models optimized based on a) winter data, b) summer data, and c) data of all four seasons of models trained with 15,799 samples, 8760 samples, and 4380 of the first and the second half and of the middle of each season of one year.



**Fig. 11.** Nowcast and global metrics for mean of all modeling targets and of all substations for the model with the lowest MSE in hyperparameter optimization for the substations with 90 % data collection rate and no detected anomalies. For the model of 192 substations, when considering those 166 without anomalous patterns, the figure depicts a) the nowcast as the mean of all targets of prediction and all substations, and the momentary error metrics b) the SE, c) the MaxAbsE, d) the RSE, and e) the AE.

comparing the AE in Fig. 11 e) to the other metrics, the difference between the heating and summer periods is less pronounced. It must be concluded that the greater sensitivity of the squared metrics and the MaxAbsE to outliers causes the pattern in Fig. 11, which suggests higher deviations in the summer. Concerning MaxAbsE (Fig. 11 b)), around 5000 h, the values exceed the maximum possible interval of one for min-max-scaled data.

Within the study's objective, the summer appears to be the most challenging period, exhibiting higher modeling errors. This is partly due to increased system dynamics in summer and the lower volume of data collected during this season, which is caused by a surplus of biomass that inhibits remote data collection and creates more extensive data gaps. In this setup, data gaps longer than three hours were filled with a constant value of  $-1$ . Consequently, the model has learned to fill in this value during summer, potentially explaining the MaxAbsE exceeding one. Therefore, data scarcity likely contributes to reduced modeling accuracy in the summer. Additionally, Idowu et al. (2014b) reports a standard deviation of  $\pm 5.71$  % MAPE for a 24-h prediction horizon during the heating period. Compared to the results of this study, which show standard deviations of  $\pm 8.87$  % (MAE) and  $\pm 22.01$  % (RMSE) relative to a maximum value of 100 %, errors are 3.16 percentage points, respectively, 16.3 percentage points higher. These differences stem particularly from the inclusion of the entire year in the evaluation, incorporating the summer's higher dynamics. Future research could focus on improving summer modeling by considering the impact of human behavior and data scarcity on model performance.

### 3.2.3. Evaluation of the heat demand

The prediction target in existing literature is the heat demand (Idowu et al., 2014b). Therefore, Fig. 12 a) shows explicitly the results for the calculated heat demand as the mean of all substations for the nowcasts.

Fig. 12 again depicts b) the SE, c) the MaxAbsE, d) the RSE, and e) the AE for the current mean across all substations.

Compared to the average results across all modeling targets, the heat demand estimates show less variability during the summer. Overall, the error metrics in Fig. 12 b)–e) indicate an increase in error during the heating season. The summer error increase for the heat demand is less significant than that for the mean of all modeling targets. This suggests that the deviations of the individual modeling targets tend to cancel out when calculating heat demand. Future research could explore whether this damping effect is consistent across different regions. As discrepancies between the modeling estimates and test data grow in summer, further work should aim to enhance the model's ability to manage randomness caused by human behavior, possibly by adding more features or using a more advanced ML architecture.

For comparison with existing literature, especially the reference value in Idowu et al. (2014b), a model with a 24-h prediction horizon is optimized using the same setup as for the nowcasts. Fig. 13 a) displays the test data, the nowcast, and the 24-h horizon prediction. Additionally, Fig. 13 compares the same parameters for an exemplary winter day (b) and summer day (c).

This figure emphasizes that the prediction model has less flexibility than the nowcast, as shown by the lower visible model variance in Fig. 13 a). When averaging across all 24-h sequences in the evaluation data, the model achieves an MAE of  $0.0612 \pm 0.0878$  and an RMSE of  $0.1070 \pm 0.2149$  for heat demand predictions. These figures indicate an increase of 0.12 percentage points in MAE and a decrease of 0.10 percentage points in RMSE compared to the nowcast. Notably, around 7000 h, the prediction significantly outperforms the nowcast, contributing to the lower RMSE in that interval. In Fig. 13 b), neither the nowcast nor the prediction perfectly matches the test data, consistently underestimating heat demand for the given day. Future studies should

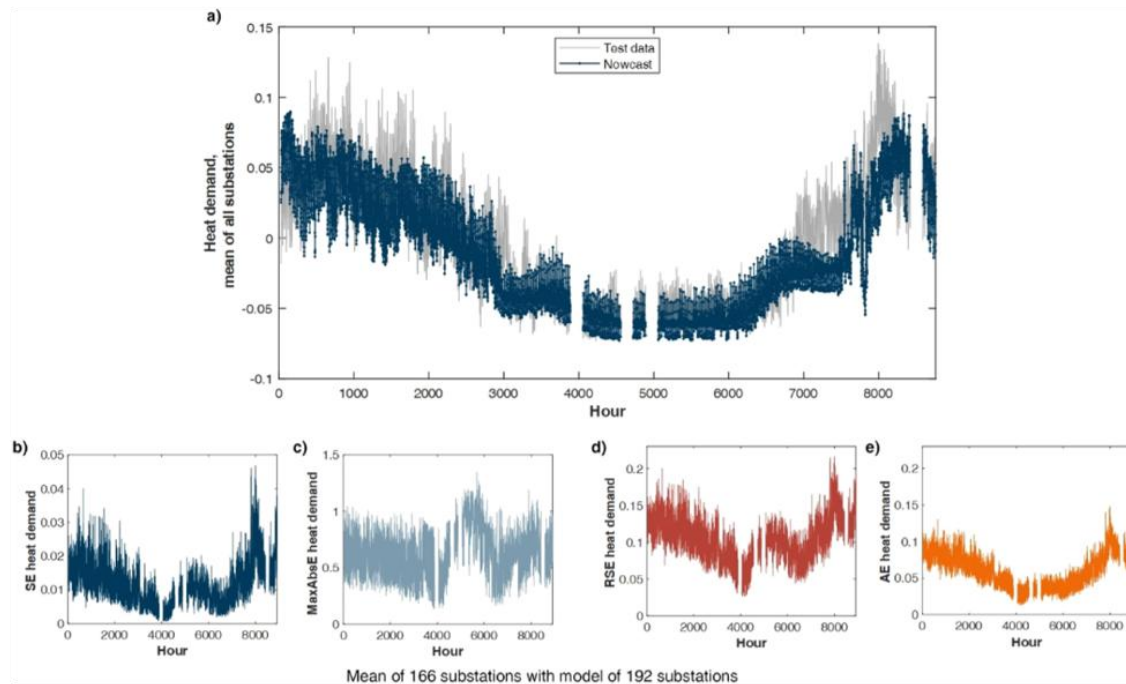


Fig. 12. Nowcast and global metrics for heat demand as mean of all substations for the model with the lowest MSE in hyperparameter optimization for the substations with 90 % data collection rate and no detected anomalies. For the model of 192 substations, when considering those 166 without anomalous patterns, the figure depicts a) the heat demand nowcast as the mean of all substations, and the momentary error metrics b) the SE, c) the MaxAbsE, d) the RSE, and e) the AE.

explore reasons for this underestimation, such as the representativeness of the training data. The prediction trend in Fig. 13 b) appears damped compared to the actual dynamics depicted by the nowcast. In Fig. 13 c), the nowcast accurately estimates the 6 am. demand peak, whereas the prediction is less precise, reaffirming its lower flexibility. Future work should focus on enhancing the model's flexibility and extending the prediction horizon. Additionally, it is essential to quantify the impact of forecasting errors inherited from operational and meteorological features. Compared to a reference MAPE of 21.5455 % (Idowu et al., 2014b), the intensive data pre-preparation, hyperparameter tuning, and domain-specific investigations on data sampling strategies in this study achieved error levels acceptable for residential settings with limited occupants. Further research could examine the damping effect observed over the prediction horizon and refine the modeling approach to understand implications for energy efficiency and system flexibility better.

#### 4. Conclusions

In the effort to transform heat networks toward climate neutrality, understanding user-level parameters is crucial for optimizing operations and enhancing efficiency. In this study, an LSTM-based modeling framework was developed for predicting user-level parameters in heat networks with high dynamics, as is the case in heat networks with comparably small user units. Focused on a northern German heat network, the framework estimates volume flow, supply, and return temperatures for residential buildings with few occupants and high data variability. Improving modeling quality through detailed analysis of

data scaling, transformation, hyperparameter tuning, and seasonal variations, even during data fluctuations, is presented, achieving error levels similar to those reported in previous studies for lower dynamics. The process began with data pre-preparation—cleaning, scaling, and transforming—using data from 15 substations supplying residential units. To validate and test scalability, models were also optimized for additional substations with at least 90 % data collection rates. This research successfully integrated an LSTM architecture with dynamic user-level heat data, laying the foundations for future exploration into modeling end-user behavior, addressing data scarcity, and testing across various datasets. Such advancements can support more reliable and efficient heat networks through improved automated ML techniques. The following findings turned out:

- Scaling non-binary, non-cyclic inputs to zero mean and unit variance improved model performance; sine-cosine transformations aided cyclic data handling. These methods reduced computational time and balanced variance and bias in modeling dynamic data.
- Hyperparameter tuning with Bayesian optimization was streamlined through interval analysis, reducing training data requirements to approximately six months to capture seasonal variations.
- The lowest MSE achieved was 0.0432, improving to 0.0392 when excluding anomalous substation data. Summer data shortages increased error margins, highlighting the importance of comprehensive data collection. Heat demand predictions showed less deviation, with MAE of 0.0600 and RMSE of 0.1071.

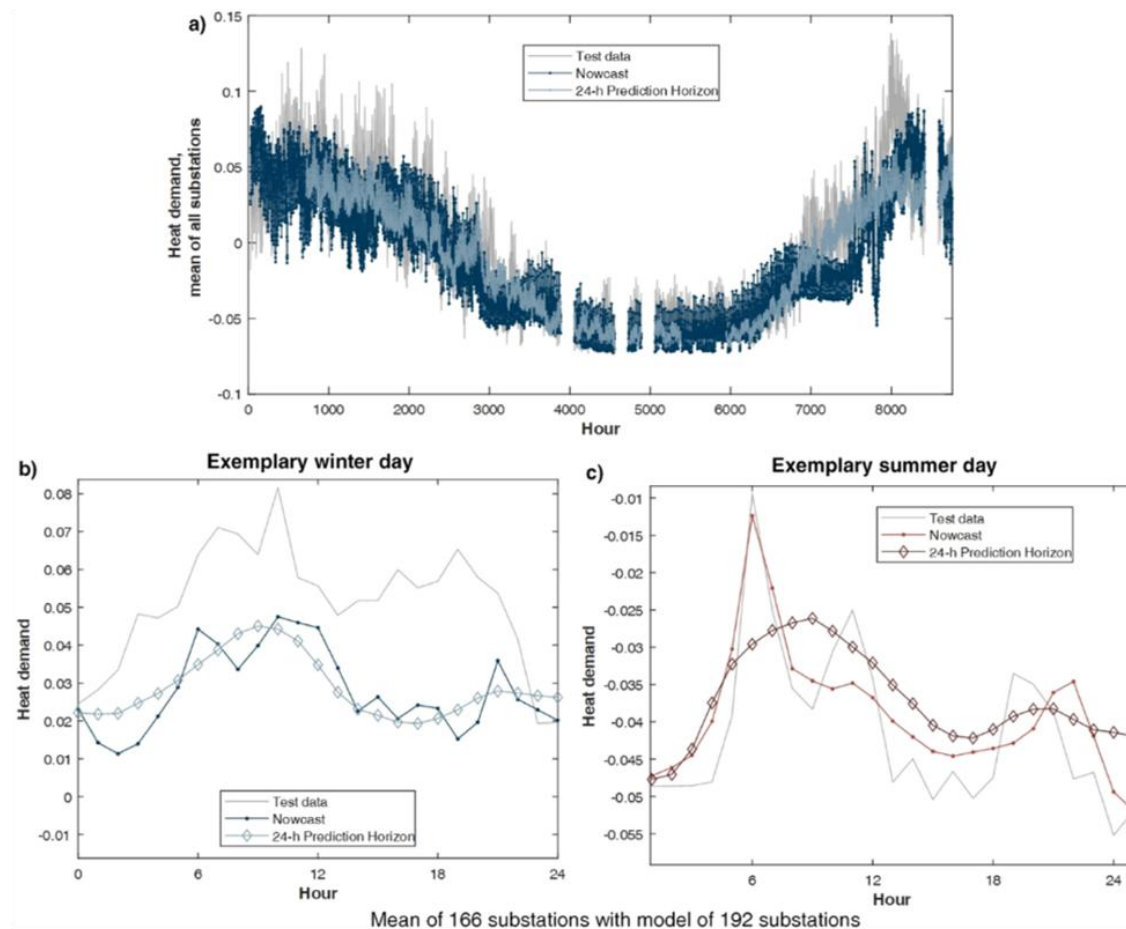


Fig. 13. Prediction with 24-h horizon, nowcast and test data for heat demand as mean of all substations for the model with the lowest MSE in hyperparameter optimization the substations with 90 % data collection rate and no detected anomalies. For the model of 192 substations, when considering those 166 without anomalous patterns, the figure depicts the heat demand prediction with a 24-h horizon, the nowcast and the test data as the mean of all substations for a) the evaluation data of 2023, b) an exemplary day of winter, c) an exemplary day of summer.

- Extending the forecast horizon to 24 h slightly reduced RMSE but raised MAE, aligning with existing benchmarks. Despite data complexities, the model attained error margins comparable to a benchmark MAPE of about 21.55 %.

#### CRedit authorship contribution statement

**Weber Samanta:** Writing – original draft, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Dirk Volta:** Supervision, Funding acquisition. **Michael Fischlschweiger:** Supervision, Investigation, Formal analysis.

#### Declaration of Generative AI and AI-assisted technologies in the writing process

Generative AI (ChatGPT (OpenAI), Grammarly (Grammarly, Inc.)) was used exclusively for linguistic correction and improvement. No AI-generated content was used in the study. After linguistic improvement, the authors reviewed and edited again and take responsibility for the content of the published article.

#### Data Availability

Raw data for the model region are not publicly available to preserve individuals' privacy under the European General Data Protection Regulation. Anonymized datasets generated for this study are available from the corresponding author upon reasonable request and with the heat network operator's consent.

#### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgments

This research was funded by the Stadtwerke Flensburg GmbH and the Society for Energy and Climate Protection Schleswig-Holstein (Gesellschaft für Energie und Klimaschutz Schleswig-Holstein GmbH, EKSH), Grant nos. 8/12-57 and 14/12-34. We would further like to

S.A. Weber et al.

Energy Reports 14 (2025) 2720–2731

express our gratitude to Stadtwerke Flensburg GmbH for providing the research data.

#### Employment

Samanta A. Weber has been employed at the Stadtwerke Flensburg GmbH, while this study was submitted for publication.

#### Additional Information

Correspondence and requests for materials should be addressed to S. W.

#### Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.egy.2025.09.032.

#### Data availability

The authors do not have permission to share data.

#### References

- Balali, Y., Chong, A., Busch, A., O'Keefe, S., 2023. Energy modelling and control of building heating and cooling systems with data-driven and hybrid models—a review. *Renew. Sustain. Energy Rev.* 183, 113496. <https://doi.org/10.1016/j.rser.2023.113496>.
- Bujalski, M., Madejski, P., Fuzowski, K., 2023. Day-ahead heat load forecasting during the off-season in the district heating system using generalized additive model. *Energy Build.* 278, 112630. <https://doi.org/10.1016/j.enbuild.2022.112630>.
- Chujai, P., Kerdprasop, N., Kerdprasop, K., 2013. Time series analysis of household electric consumption with ARIMA and ARMA models. In: *Proceedings of the International MultiConference of Engineers and Computer Scientists 2013 Vol I*, (IMECS 2013), March 13–15. Hong Kong.
- Čirić, I., et al., 2020. Intelligent heat demand prediction for advanced district heat plant control. In: *Zdravković, M., Konjović, Z., Trajanović, M. (Eds.), ICIST 2020 Proceedings*, pp. 198–201.
- Greff, K., Srivastava, R.K., Koutník, J., Steunebrink, B.R., Schmidhuber, J., 2017. LSTM: a search space odyssey. *IEEE Trans. Neural Netw. Learn. Syst.* 28 (10), 2222–2232. <https://doi.org/10.1109/TNNLS.2016.2582924>.
- Hou, H., et al., 2022. Review of load forecasting based on AI methodologies, models, and challenges. *Electr. Power Syst. Res.* 210, 108067. <https://doi.org/10.1016/j.epsr.2022.108067>.
- Houran, M.A., Bukhari, S.M.S., Zafar, M.H., Mansoor, M., Chen, W., 2023. COA-CNN-LSTM: Coati optimization algorithm-based hybrid deep learning model for PV/wind power forecasting in smart grid applications. *Appl. Energy* 349, 121638. <https://doi.org/10.1016/j.apenergy.2023.121638>.
- Idowu, S., Åhlund, C., Schelén, O., 2014a. Machine learning in district heating system energy optimization. 2014 IEEE International Conference on Pervasive Computing and Communication Workshops. PERCOM WORKSHOPS, Budapest, Hungary, pp. 224–227. <https://doi.org/10.1109/PerComW.2014.6815206>.
- Idowu, S., Saguna, S., Åhlund, C., Schelén, O., 2014b. Forecasting heat load for smart district heating systems: a machine learning approach. In: *Proceedings of the 2014 IEEE International Conference on Smart Grid Communications (SmartGridComm)*. Venice, Italy, pp. 554–9. <http://doi.org/10.1109/SmartGridComm.2014.7007705>.
- Idowu, S., Saguna, S., Åhlund, C., Schelén, O., 2016. Applied machine learning: forecasting heat load in district heating system. *Energy Build.* 133, 478–488. <https://doi.org/10.1016/j.enbuild.2016.09.068>.
- Index of climate\_environment/CDC/observations\_germany/climate/hourly/. ([https://opendata.dwd.de/climate\\_environment/CDC/observations\\_germany/climate/hourly/](https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/hourly/)), (Accessed 14 July 2023).
- Jiang, B., Gong, H., Qin, H., Zhu, M., 2022. Attention-LSTM architecture combined with Bayesian hyperparameter optimization for indoor temperature prediction. *Build. Environ.* 224, 109536. <https://doi.org/10.1016/j.buildenv.2022.109536>.
- Johansson, C., et al., 2017. Operational demand forecasting in district heating systems using ensembles of online machine learning algorithms. *Energy Procedia* 116, 208–216. <https://doi.org/10.1016/j.egypro.2017.05.068>.
- Kemper, N., Heider, M., Pietruschka, D., Hähner, J., 2025. Forecasting of residential unit's heat demands: a comparison of machine learning techniques in a real-world case study. *Energy Syst.* 16, 281–315. <https://doi.org/10.1007/s12667-023-00579-y>.
- Kim, J., Kim, S., Choi, S., 2018. Learning to Warm-Start Bayesian Hyperparameter Optimization. arXiv:1710.06219, Version 3. (<https://doi.org/10.48550/arXiv.1710.06219>).
- Laakkonen, L., et al., 2017. Predictive supply temperature optimization of district heating networks using delay distributions. *Energy Procedia* 116, 297–309. <https://doi.org/10.1016/j.egypro.2017.05.076>.
- Leiprecht, S., Behrens, F., Faber, T., Finkenrath, M., 2021. A comprehensive thermal load forecasting analysis based on ML algorithms. *Energy Rep.* 7 (Suppl. 4), S319–S326. <https://doi.org/10.1016/j.egy.2021.08.140>.
- Mahajan, T., et al., 2021. An experimental assessment of treatments for cyclical data. In: *Proceedings of the 2021 Computer Science Conference for CSU Undergraduates*, Virtual, 6, p. 22. (<http://hdl.handle.net/20.500.12680/th83m446n>).
- MathWorks: Choose training configuration for LSTM using Bayesian Optimization – MATLAB & Simulink - MathWorks Deutschland. Available online: (<https://de.mathworks.com/help/deeplearning/ug/exp-mgr-sequence-regression-example.html#SequenceRegressionExample-8>). (Accessed 3rd August 2023).
- MathWorks: Options for training deep learning neural network - Matlab trainingOptions - MathWorks Deutschland. Available online: ([https://de.mathworks.com/help/deeplearning/ref/trainingoptions.html#bu59f0g\\_sep\\_mw\\_da5dd7aa-a922-4f38-b162-52f5115795f6\\_head](https://de.mathworks.com/help/deeplearning/ref/trainingoptions.html#bu59f0g_sep_mw_da5dd7aa-a922-4f38-b162-52f5115795f6_head)). (Accessed 30th January 2024).
- Mbiydenyuy, G., et al., 2021. Opportunities for machine learning in district heating. *Appl. Sci.* 11, 6112. <https://doi.org/10.3390/app11136112>.
- Ntakolia, C., Anagnostis, A., Moustakidis, S., Karcianias, N., 2022. ML applied on the district heating and cooling sector: a review. *Energy Syst.* 13, 1–30. <https://doi.org/10.1007/s12667-020-00405-9>.
- Potočník, P., Škerl, P., Govekar, E., 2021. Machine-learning-based multi-step heat demand forecasting in a district heating system. *Energy Build.* 233, 110673. <https://doi.org/10.1016/j.enbuild.2020.110673>.
- Singh, D., Birmohan Singh, B., 2020. Investigating the impact of data normalization on classification performance. *Appl. Soft Comput.* 97 (part B), 105524. <https://doi.org/10.1016/j.asoc.2019.105524>.
- Snoek, J., Larochelle, H., Adams, R.P., 2012. Practical Bayesian Optimization of Machine Learning Algorithms. arXiv:1206.2944 (stat.ML), Version 2, pp. 1–12. (<https://doi.org/10.48550/arXiv.1206.2944>).
- Song, J., et al., 2021. Predicting hourly heating load in a district heating system based on a hybrid CNN-LSTM model. *Energy Build.* 243, 110998. <https://doi.org/10.1016/j.enbuild.2021.110998>.
- Technische Anschlussbedingungen (TAB) für den Anschluss an das Fernwärmenetz Tarp der Stadtwerke Flensburg GmbH, 2011. ([https://www.stadtwerke-flensburg.de/fileadmin/user\\_upload/PDFs/1.2\\_Service/1.2.7\\_Downloads/Fernwaerme/06\\_-Technische\\_Anschlussbedingungen.pdf](https://www.stadtwerke-flensburg.de/fileadmin/user_upload/PDFs/1.2_Service/1.2.7_Downloads/Fernwaerme/06_-Technische_Anschlussbedingungen.pdf)), (Accessed 11th April 2024).
- Weber, S.A., Fischlschweiger, M., Volta, D., Geisler, J., 2025. Feature selection for specific prediction targets at the user level in a district heating network. *Sci Rep* 15, 29789. <https://doi.org/10.1038/s41598-025-15777-0>.
- Weber, S., Fischlschweiger, M., Volta, D., Rieck-Blankenburg, U., 2024. Clustering- and statistic-based approach for detection and impact evaluation of faults in end-user substations of thermal energy systems. *Sci. Rep.* 14, 32166. <https://doi.org/10.1038/s41598-024-82103-5>.

### 3.3.3 Supplemental Material of Publication 3

Supplementary Material

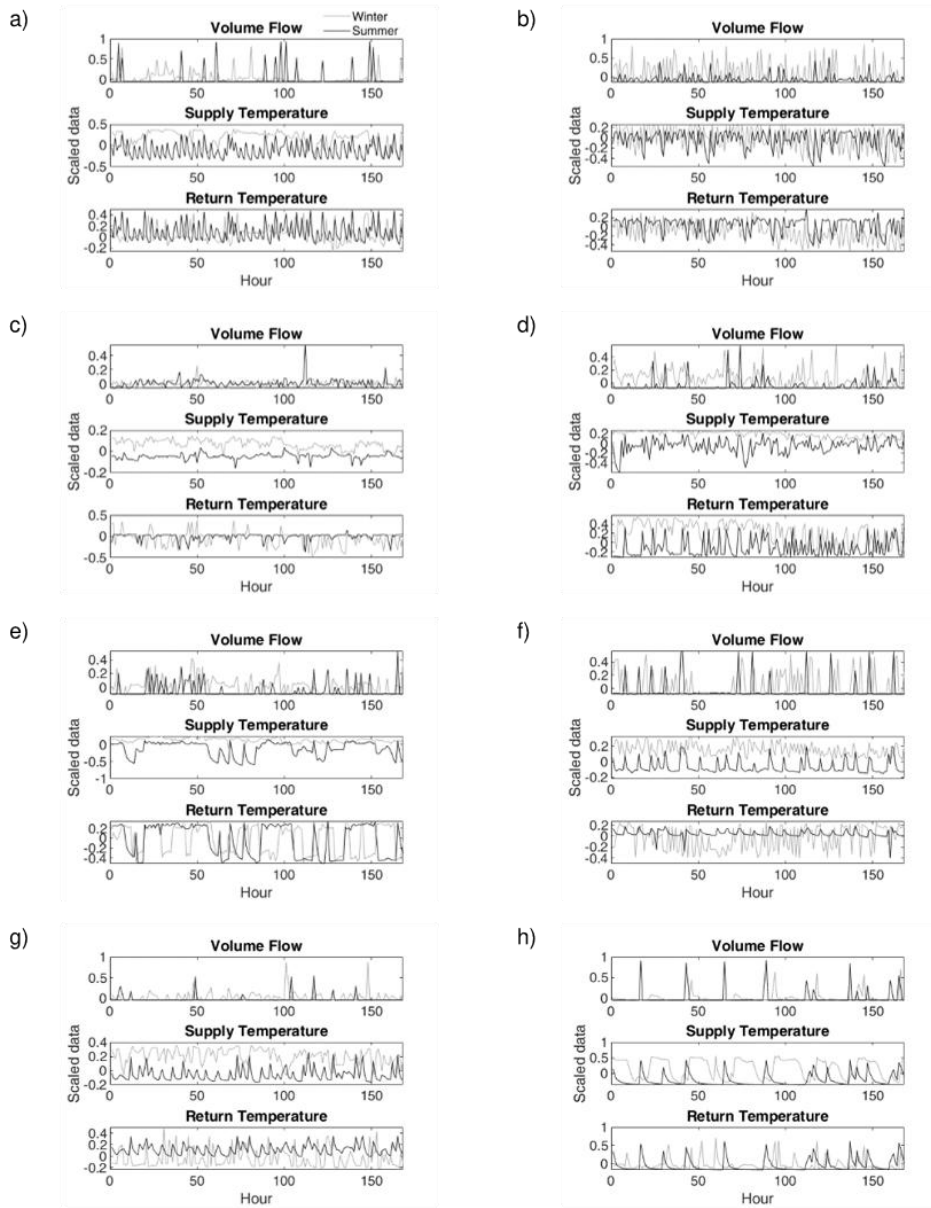
## Dynamic user-level parameter modeling in heat networks with deep learning Long Short-Term Memory

Samanta A. Weber<sup>a,b,\*</sup>, Michael Fischlschweiger<sup>b</sup> & Dirk Volta<sup>a</sup>

<sup>a</sup> Energy and Life Science, University of Applied Sciences Flensburg, 24943 Flensburg, Germany

<sup>b</sup> Chair of Technical Thermodynamics and Energy Efficient Material Treatment, Institute for Energy Process Engineering and Fuel Technology, Clausthal University of Technology, 38678 Clausthal-Zellerfeld, Germany

\* Corresponding author: [samanta.weber@tu-clausthal.de](mailto:samanta.weber@tu-clausthal.de), [samanta.weber@hs-flensburg.de](mailto:samanta.weber@hs-flensburg.de), <https://orcid.org/0000-0002-6487-8248>



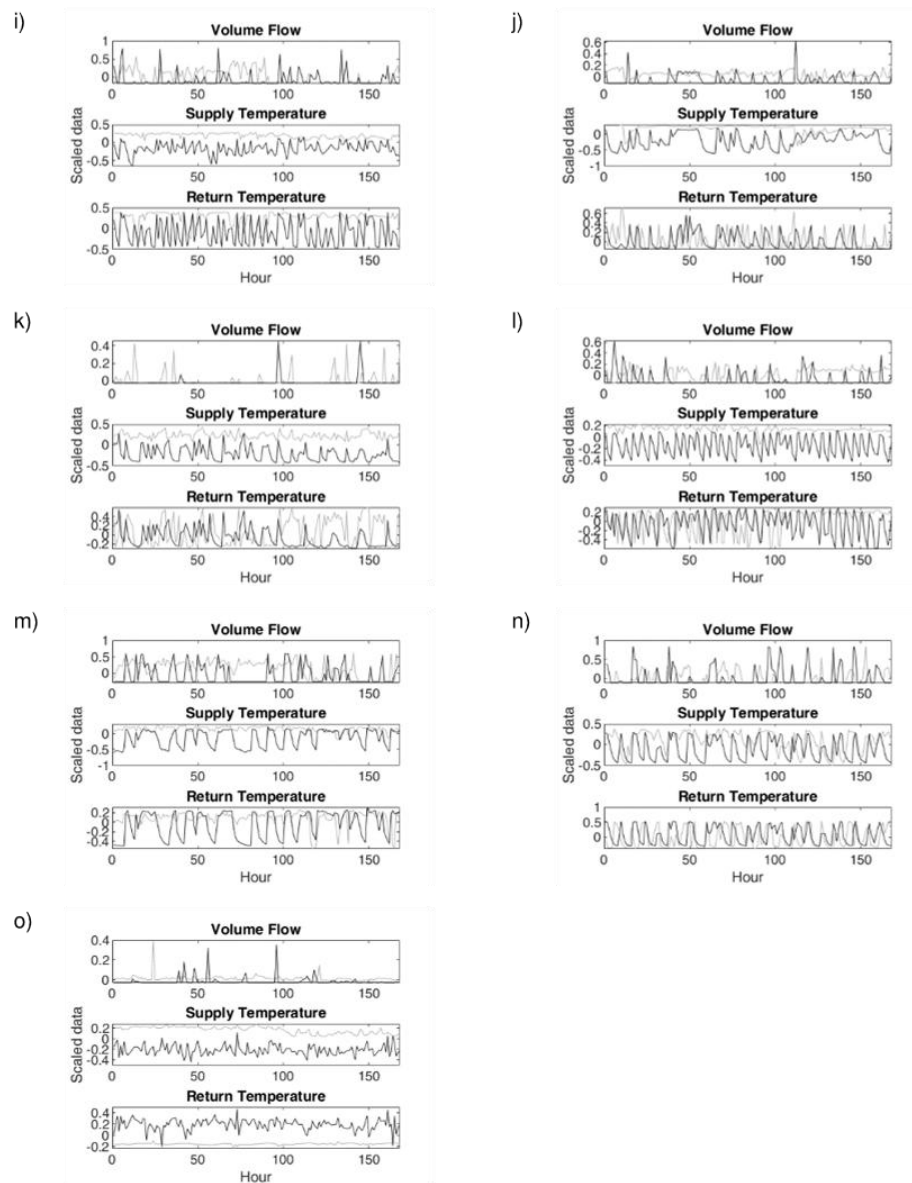


Figure 3 Exemplary data for the 15 substations investigated, for a week in winter and in summer. The figure captures the modeling targets individually and pinpoints the intensity of the dynamics as well as the change in temperature level and the apparent randomness between the seasons.

Table 2 Hyperparameter ranges for optimization given in literature and chosen settings of this study

| Hyperparameter                                       | Ranges in Literature                                          | Chosen range                        | Type/Transformation method |
|------------------------------------------------------|---------------------------------------------------------------|-------------------------------------|----------------------------|
| Initial learning rate                                | *[0.0001, 0.0005, 0.001, 0.002, 0.005]<br>*[0.001, 0.1]       | [0.0001, 0.02]                      | real/log                   |
| Dropout rate                                         | **[0.1, 0.2]<br>*[0.5]                                        | [0.1, 0.5]                          | real                       |
| Validation frequency                                 | *[30]                                                         | [10, 100]                           | integer                    |
| Validation patience factor                           | -                                                             | [1, 10]                             | integer                    |
| LSTM-depth                                           | *[1,3]                                                        | [1, 5]                              | integer                    |
| Number of hidden units (LSTM-layer)                  | *[10,12,15,20,25]<br>**[5,8,10,12,15]<br>*[50,300]            | [10, 300]                           | integer                    |
| Number of hidden units (first fully connected layer) | *[10]<br>*[100]                                               | [10, 300]                           | integer                    |
| Mini-batch size                                      | *[5, 10, 15, 20, 25]<br>** [3, 4, 5, 6, 7, 8, 9, 10]<br>*[20] | [1, 10]                             | integer                    |
| Sequence length <sup>1</sup>                         | No reference                                                  | [1,2,6, 12,24, 168,720, 4380, 8760] | categorical                |
| Solver                                               | *[sgdm,adam]<br>**[sgdm,adam]                                 | [adam, sgdm, rmsprop]               | categorical                |
| Training data partition                              | No reference                                                  | [0.5, 0.9]                          | real                       |

|                              |                                                                                                                |                                |             |
|------------------------------|----------------------------------------------------------------------------------------------------------------|--------------------------------|-------------|
| Iterations/Epochs            | *[50, 150, 200, 250, 300]<br>(iterations)<br><br>** [30, 50, 70, 90, 110]<br>(epochs)<br><br>'max 300 (epochs) | Max 300<br>(epochs)            | integer     |
| Learning rate drop factor    | '[0.2]<br><br>''[0.1]                                                                                          | [0.2]                          | real        |
| Learning rate drop period    | '[15]<br><br>''[10]                                                                                            | [15]                           | integer     |
| Output Network               | ''["last iteration", "best-validation-loss"]                                                                   | "best-validation-loss"         | categorical |
| Shuffle                      | ''["once", "every epoch", "never"]                                                                             | „never“<br><br>(done manually) | categorical |
| *[17], **[20], '[15], ''[25] |                                                                                                                |                                |             |

<sup>1</sup>The “SequenceLength” property of Matlab® either increases the length of all sequences in a batch to the longest (default and applied here) or reduces it to the shortest.

Table 3 Cleaning methods applied to the operational data

| Parameter at CHP              | Treatment of outliers                                                                                                                                         |
|-------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Supply temperature            | Previous value using moving mean over 30-day window, threshold factor 5.5; Previous value using moving mean over 14-day window, threshold factor 3.5          |
| Return temperature            | Previous value using moving mean over 30-day window, threshold factor 2.75                                                                                    |
| Volume Flow                   | Linear interpolation using mean; Previous value using mean; Previous value using moving median over 30-day window, threshold factor 3.25                      |
| Accumulated heat (no feature) | Linear interpolation using mean; Next value using moving mean over 7-day window; Linear interpolation using moving mean over 3-day window, threshold factor 2 |
| Heat flow                     | Linear interpolation using mean, threshold factor 3.75; Previous value using moving median over 30-day window                                                 |

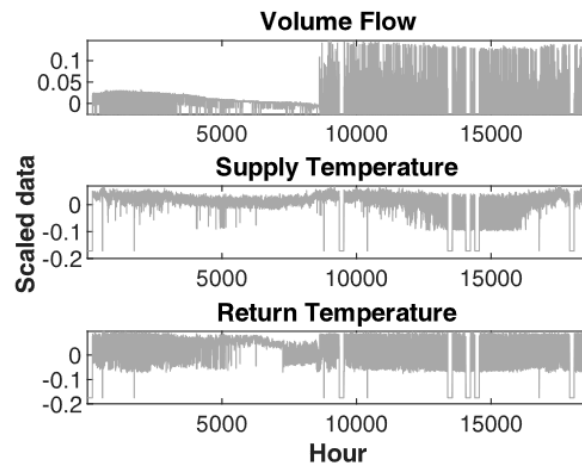
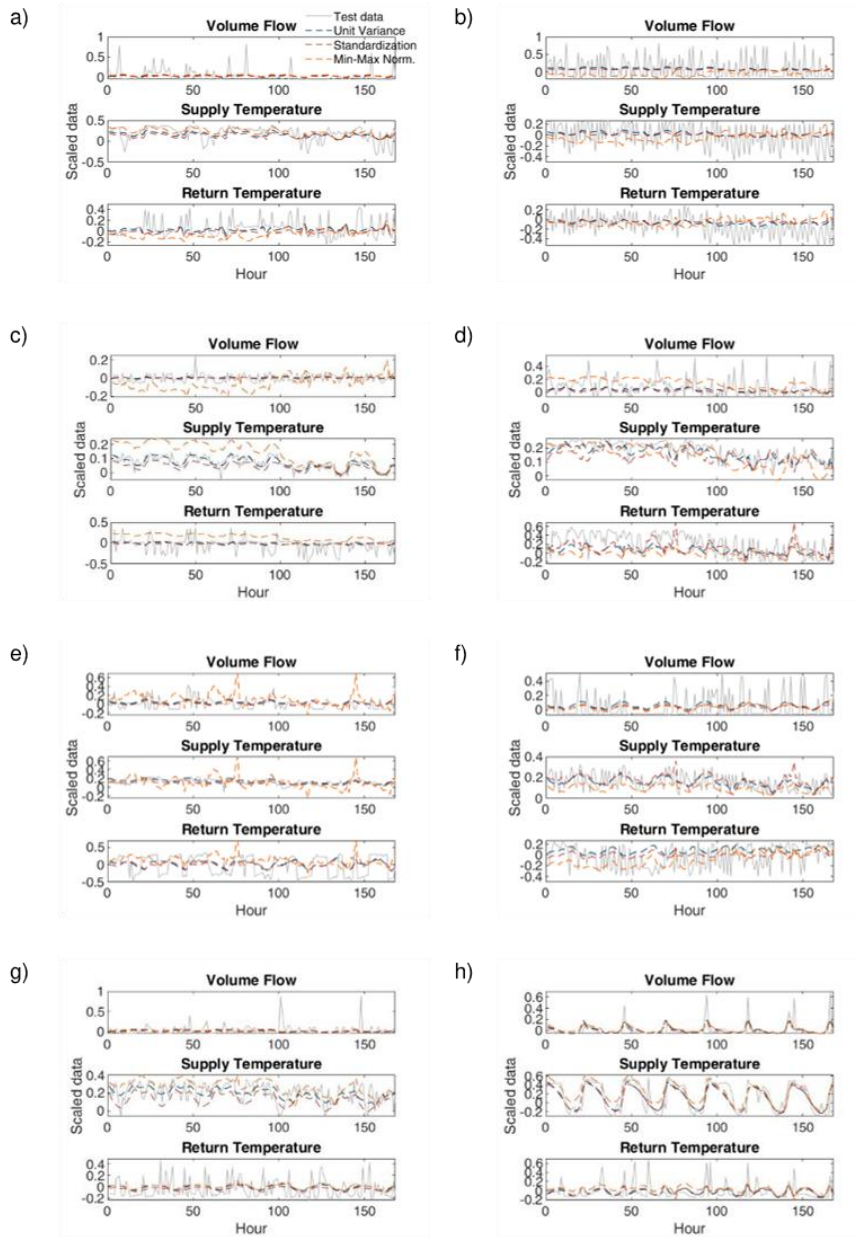


Figure 4 Time series data for an exemplary substation. The figure suggests a significant change in behavior during the measurement period, especially within the volume flow data.



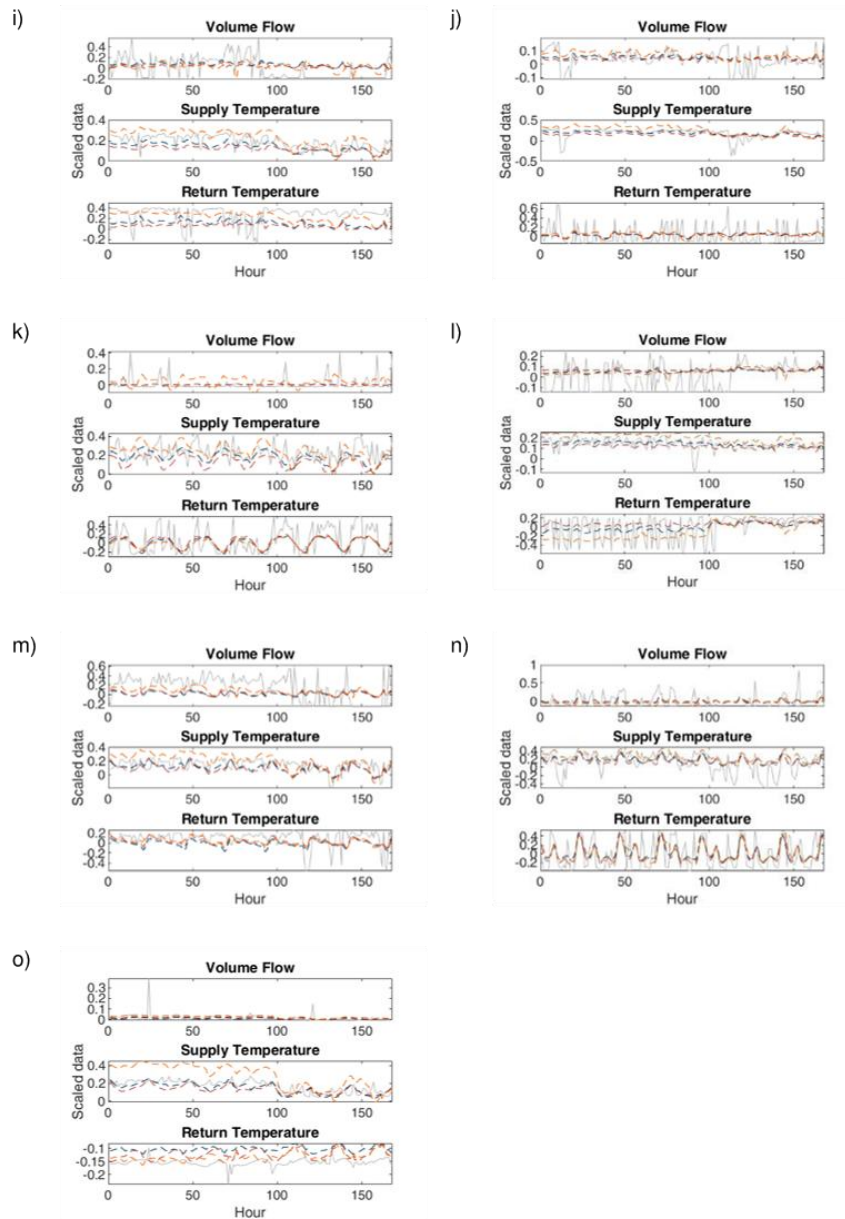


Figure 6 Nowcasts of optimized models for 15 exemplary substations. The models were trained using unit variance (dark blue line), standardization (red line), and min-max normalization (orange line) for scaling, and their performance was compared to the test data (grey line).

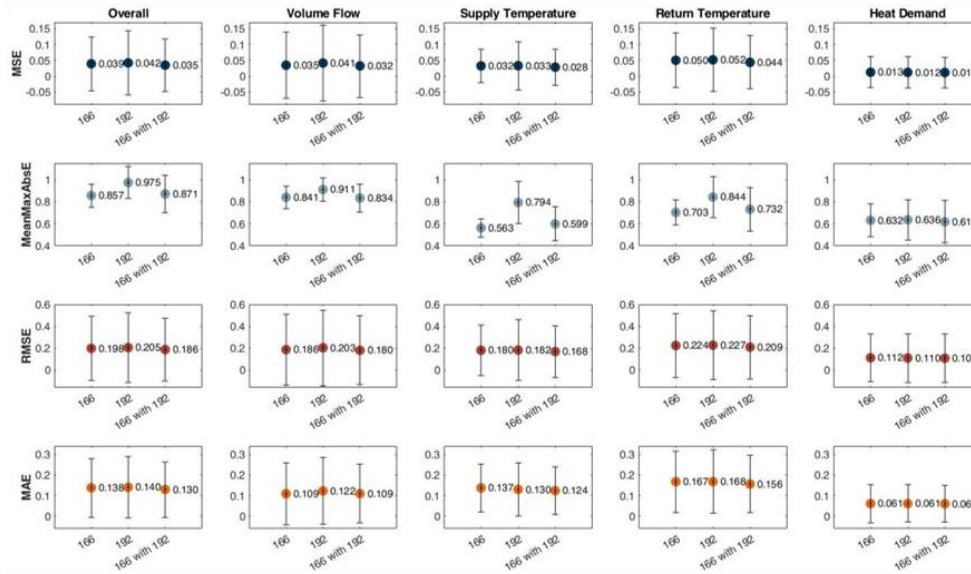


Figure 10 Visualization of results for global model metrics for evaluation data of 2023. The figure visualizes the metric values and standard deviations for the mean of all targets and substations, each prediction target separately, and the calculated heat demand for the models of 166 and 192 substations and the model of 92 substations when excluding those with anomalous patterns referenced as “166 with 92”.

Table 4 Results for global model metrics for evaluation data of 2023. The background coloring highlights the higher values in darker colors for each metric, the frames outline the values for comparison to the literature reference, the overall values are in bold.

| Model →<br>Metric ↓ | Model for 166<br>substations |                | Model for 192<br>substations |                | Model for 192,<br>results of 166<br>substations |                |
|---------------------|------------------------------|----------------|------------------------------|----------------|-------------------------------------------------|----------------|
| <b>MSE</b>          | <b>0.0390</b>                | <b>±0.0845</b> | <b>0.0420</b>                | <b>±0.1005</b> | <b>0.0348</b>                                   | <b>±0.0819</b> |
| Volume flow         | 0.0347                       | ±0.1049        | 0.0414                       | ±0.1198        | 0.0323                                          | ±0.0990        |
| Supply temperature  | 0.0324                       | ±0.0536        | 0.0331                       | ±0.0762        | 0.0284                                          | ±0.0561        |
| Return temperature  | 0.0501                       | ±0.0858        | 0.0516                       | ±0.0999        | 0.0436                                          | ±0.0840        |
| Heat flow           | 0.0125                       | ±0.0484        | 0.0117                       | ±0.0447        | 0.0115                                          | ±0.0484        |
| <b>MeanMaxAbsE</b>  | <b>0.8568</b>                | <b>±0.1044</b> | <b>0.9750</b>                | <b>±0.1462</b> | <b>0.8712</b>                                   | <b>±0.1688</b> |
| Volume flow         | 0.8409                       | ±0.1022        | 0.9115                       | ±0.1063        | 0.8344                                          | ±0.1246        |
| Supply temperature  | 0.5627                       | ±0.0828        | 0.7937                       | ±0.1932        | 0.5990                                          | ±0.1554        |
| Return temperature  | 0.7032                       | ±0.1152        | 0.8435                       | ±0.1883        | 0.7318                                          | ±0.1999        |
| Heat flow           | 0.6317                       | ±0.1496        | 0.5885                       | ±0.1611        | 0.6179                                          | ±0.1928        |
| <b>RMSE</b>         | <b>0.1976</b>                | <b>±0.2907</b> | <b>0.2050</b>                | <b>±0.3170</b> | <b>0.1864</b>                                   | <b>±0.2862</b> |
| Volume flow         | 0.1863                       | ±0.3239        | 0.2034                       | ±0.3461        | 0.1798                                          | ±0.3146        |
| Supply temperature  | 0.1799                       | ±0.2314        | 0.1820                       | ±0.2761        | 0.1685                                          | ±0.2368        |
| Return temperature  | 0.2237                       | ±0.2929        | 0.2272                       | ±0.3160        | 0.2087                                          | ±0.2899        |
| Heat flow           | 0.1118                       | ±0.2200        | 0.1080                       | ±0.2115        | 0.1071                                          | ±0.2201        |
| <b>MAE</b>          | <b>0.1377</b>                | <b>±0.1417</b> | <b>0.1401</b>                | <b>±0.1497</b> | <b>0.1296</b>                                   | <b>±0.1340</b> |
| Volume flow         | 0.1090                       | ±0.1511        | 0.1224                       | ±0.1624        | 0.1093                                          | ±0.1428        |
| Supply temperature  | 0.1368                       | ±0.1168        | 0.1300                       | ±0.1273        | 0.1236                                          | ±0.1144        |
| Return temperature  | 0.1674                       | ±0.1484        | 0.1678                       | ±0.1531        | 0.1559                                          | ±0.1387        |
| Heat flow           | 0.0610                       | ±0.0937        | 0.0610                       | ±0.0892        | 0.0600                                          | ±0.0887        |

## 4 Discussion, Conclusions, and Outlook

Following the objective of making information on heat network efficiency and flexibility accessible through modeling and analysis of building-level data that reflect highly dynamic user-level behavior, this cumulative dissertation contributes to the literature through a three-subobjective research design. To enable a holistic evaluation of these contributions, the following sections synthesize and discuss the investigations presented, conclude, and provide an outlook to summarize this cumulative dissertation.

Starting with the discussion, while each publication separately discusses its findings, the cumulative perspective adopted here provides an overarching interpretation and contextualization of the results.

Central to all investigations is the data upon which the models and analyses are based. The depth and width of the inquiries in this cumulative dissertation depended, in certain respects, on data availability, as shall be explained below. Overall, data scarcity was a relevant boundary condition in this cumulative dissertation. Thus, a significant objective for the future is to make user-level heat network data widely available for research, e.g., through data platforms.

Collectively, this cumulative dissertation demonstrates the successful implementation and optimization of concepts that pursue the overarching research objective for the investigated model region. The concepts were tested across the model region and scaled up to as many substations as provided sufficient data. The promising results motivate further inquiry into the framework's transferability across different heat networks, which have remained constrained by data scarcity, particularly regarding the extent to which model development and optimization can be automated.

To ensure practical feasibility and real-world applicability, this cumulative dissertation relies on readily available measured data. To evaluate the interdependencies between the model and manually selected settings, the developed methodology employs sensitivity and comparative analyses. The findings demonstrate a reduction in modeling error through thorough optimization of the modeling framework. Likewise, subsequent inquiries can explore additional parameters as features in ML setups, propose alternative measurements for model building, and assess whether the benefits justify the investment in integration.

For instance, in the context of data, this cumulative dissertation examines the significance of data quality. This relates, e.g., to large measurement gaps, anomalous behavior, or erroneous readings. Excluding substations exhibiting anomalous behavior significantly improved modeling performance. Regarding the importance of high-quality data, this cumulative dissertation presents methods to enhance data quality (e.g., filling missing values, reducing noise, and discarding outliers). To progress further, the existing measurement error could be addressed in terms of its impact on ML modeling performance. Evaluating which measurement accuracy best serves the objective of user-level modeling could prove to be promising. This would help determine the necessary data quality and limit the required data quantity. Moreover, this cumulative dissertation demonstrates the successful implementation of modeling frameworks and stepwise methods applied to hourly-resolution data. Subsequent research could investigate the collection at higher sampling frequencies (e.g., every 15 minutes, as in electrical grids) to determine the optimal approach for capturing the dynamics.

In particular, this cumulative dissertation addresses modeling heat-network user-level parameters for buildings with a limited number of regular occupants. As noted earlier, a key contributor to the apparent randomness and modeling complexity of such data is human behavior. Therefore, subsequent investigations aiming to advance in modeling this influencing factor could include, for example, enhanced occupation modeling. Within the context of human behavior, another area of interest is the effect of factors not directly captured by the features but indirectly reflected in the user-level data

collected. To integrate information into the model, this cumulative dissertation presents an approach that uses operational, meteorological, and temporal features. Yet, beyond the information contained in the features, the modeling results will also be influenced by human behavior in response to local, regional, or even global factors captured in the data. Thus, subsequent inquiries could examine the relevance of specific features for user-level modeling targets that reflect the effects of human behavior on heat demand at a meta-level, such as the impacts of climate, crises, and politics. This area of interest is particularly relevant to the aim of understandable Artificial Intelligence. Explicitly linking the features and settings to the model's behavior could increase the acceptance of ML applications.

Continuing with the conclusions of this cumulative dissertation, the following revisits the three subobjectives formulated to pursue the overarching objective.

- Publication 1 aimed to enable the automated detection of anomalous, especially fault-induced, substation data patterns. It demonstrates the successful application of substation temperature data for user-level fault detection in heat networks, aiming to achieve a high degree of automation. Furthermore, the proposed chain of methods enables quantification of the negative impact of faults by combining temperature information and volume-flow recordings. The findings confirm the successful identification of faults that have a relevant impact on the system. Fault elimination can significantly reduce system temperature and increase capacity, thereby enhancing efficiency and flexibility. Within the research design, the findings from pursuing subobjective 1 were essential to enhance data quality for achieving subobjective 3.
- Publication 2 explores feature engineering and selection for modeling the user-level parameters of the supply and the return temperature, as well as the volume flow, at buildings' substations. While temporal features and the operational parameters of the heat infeed facility significantly impact the target parameters, meteorological data appear to be less essential. The choice of modeling architecture and the selection of relevant, non-redundant features are interdependent, as demonstrated by comparing different feature investigation and selection algorithms. Regarding the research design of this cumulative dissertation, the inquiry into subobjective 2 provided the required information on the selection of features for pursuing subobjective 3.
- Publication 3 investigates the additional steps necessary to develop a modeling framework for the suggested user-level parameters that reflect the highly dynamic user-level. The concept using a basic LSTM architecture includes data pre-preparation, hyperparameter optimization, and domain-specific impacts of data quantity and seasonality. Based on data from a subset of characteristic substations and the features selected in publication 2, suitable data pre-preparation methods are identified, including zero-mean unit-variance scaling and sine- and cosine-based transformations of cyclic continuous features. Adequate hyperparameter ranges are identified and validated. The minimum required data quantity and the impact of seasonality are assessed. The findings demonstrate a successful upscaling across the heat network. The methods from publication 1 were used to discard substation data unsuitable for modeling. Thus, the modeling error could be reduced to a level comparable to existing benchmarks, despite the high dynamics of user-level behavior captured in the modeling targets. The information made accessible by models comparable to those established in publication 3 can bridge the information gap between operators and users. Such optimized heat network control has the potential to increase the efficiency and flexibility of heat networks, consistent with the overarching research objective of this cumulative dissertation.

Figure 8 summarizes the subobjectives, data, methods, and contributions of each publication associated with this cumulative dissertation.

| Data-Driven Modeling and Analysis of the Highly Dynamic User-Level Behavior in Heat Networks to Enhance the Efficiency and Flexibility   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Clustering- and statistic-based approach for detection and impact evaluation of faults in end-user substations of thermal energy systems | <b>Subobjective:</b><br>Dismissing datasets unsuitable for modeling with fault management approach for user-level substations with high degree of automation addressing requirements for practical feasibility.                                                                                                                                                                                                                                                                                                                                  |
|                                                                                                                                          | <b>Data:</b><br>Peer-reviewed studies from 2004-2023 on the subject of data-driven fault detection and management in heat networks, user-level data of the individual substations in the model region of 2023.                                                                                                                                                                                                                                                                                                                                   |
|                                                                                                                                          | <b>Methods:</b><br>Literature research, k-means clustering of temperature signatures transformed into heat maps, expert knowledge and labeling, statistical analysis, thermodynamic fundamentals, encoding for automation.                                                                                                                                                                                                                                                                                                                       |
|                                                                                                                                          | <b>Contributions:</b><br>Systematic three-step method for identification of indicators for the occurrence of faults in user substations, detection of faulty user substations and determination of the detected faults' negative impact on the system with high degree of automation, gain of knowledge for applying user-level temperature data not structured as a time series for fault management on the user level in heat networks including labeling data; automated filtering of datasets unsuitable for modeling within subobjective 3. |
| Feature selection for specific prediction targets at the user level in a district heating network                                        | <b>Subobjective:</b><br>Obtaining a set of relevant, non-redundant features for the modeling targets supply temperature, return temperature and volume flow at individual user substations in the model region.                                                                                                                                                                                                                                                                                                                                  |
|                                                                                                                                          | <b>Data:</b><br>Peer-reviewed studies from 2017-2023 on the subject of data-driven modeling in heat networks with special regard to feature engineering and selection, meteorological data of the DWD, operational data of the CHP and user-level data of the individual substations in the model region of 1st December 2021 to 17th June 2023.                                                                                                                                                                                                 |
|                                                                                                                                          | <b>Methods:</b><br>Literature research, domain knowledge, statistical analysis, correlation analysis, feature investigation and selection methods (RNN), combination of information on correlation and feature relevance.                                                                                                                                                                                                                                                                                                                        |
|                                                                                                                                          | <b>Contributions:</b><br>Systematic approach to feature engineering and selection for modeling parameters of the highly dynamic user level in heat networks in a chain of methods, gain of knowledge for accelerating ML modeling; set of features for user-level modeling in the model region for subobjective 3.                                                                                                                                                                                                                               |
| Dynamic user-level parameter modeling in heat networks with deep learning Long Short-Term Memory                                         | <b>Subobjective:</b><br>Development of a modeling framework for the suggested parameters of the highly dynamic user-level in heat networks and reduction of the modeling error to an acceptable level by thoroughly optimizing the modeling setup.                                                                                                                                                                                                                                                                                               |
|                                                                                                                                          | <b>Data:</b><br>Peer-reviewed studies from 2012-2024 on the subject of data-driven modeling in heat networks including data pre-preparation and hyperparameter optimization, meteorological data of the DWD, operational data of the CHP and user-level data of the individual substations in the model region of 1st December 2021 to 15th January 2024.                                                                                                                                                                                        |
|                                                                                                                                          | <b>Methods:</b><br>Literature research, Bayesian Optimization, DNN (LSTM), sensitivity and comparative analyses, scaling, sine-cosine transformation, statistical analysis, metric functions (e.g., MSE, MeanMaxAbsE).                                                                                                                                                                                                                                                                                                                           |
|                                                                                                                                          | <b>Contributions:</b><br>Basic LSTM modeling framework with systematic stepwise approach for determination of methods for data pre-preparation and hyperparameter optimization, impact assessment of data quantity and seasonality within the sampling strategy for automated model building for highly dynamic user-level behavior in heat networks, knowledge gain on the scalability of the results within the heat network and overall model performance compared to existing benchmark.                                                     |

Figure 3 Structure of this research with subobjectives, data, methods, and contributions

Lastly, closing with an outlook, the findings of this cumulative dissertation open several directions for future research that extend beyond the scope of the presented investigations while building directly on their methodological and conceptual foundations. In particular, the results highlight the value of highly dynamic user-level information for both heat network analysis and operation, suggesting directions for future research.

While this cumulative dissertation primarily focuses on analysis and knowledge generation, a conclusive subsequent step is to integrate user-level insights into operational and control strategies. Future research may explore how modeling and analysis results can serve as indicators of efficiency and flexibility, and how they can be incorporated into predictive or adaptive control frameworks. This could enable bridging the gap between offline analysis and real-time decision-making. The increasing digitalization of heat networks also provides opportunities for further development. Future work could investigate automated model calibration, online learning, and continuous updating of user-level behavior models.

Beyond operational aspects, the insights into highly dynamic user-level behavior may support the activation of demand-side flexibility. Future studies could examine the coordinated use of building thermal mass, decentralized storage, or hybrid heating systems to exploit efficiency and flexibility potentials at the user level. In combination with sector-coupling options, such as power-to-heat technologies, this could enhance the integration of renewable energy sources and increase system-wide efficiency and flexibility. Furthermore, extending the analysis to include economic and environmental dimensions offers a broader perspective. Linking user-level dynamics, for example, to costs and emissions could help quantify the general benefits of reduced network temperatures and improved operational strategies.

From a methodological perspective, future work could extend the proposed modeling and analysis approaches to explicitly account for additional sources of uncertainty. This includes uncertainties arising from weather forecasts, measurement errors, and incomplete or heterogeneous datasets. Combining the presented data-driven methods with physics-based network models may further enhance robustness and interpretability. With respect to the scalability and transferability of user-level modeling approaches, applying the methods developed in this cumulative dissertation to larger or more heterogeneous heat networks across different climatic regions, building stock, and network topologies would enable an assessment of their generalizability. Such studies could help distinguish between network-specific characteristics and more universal patterns in user behavior, thereby supporting the broader applicability of the results.

In summary, the investigations conducted in this cumulative dissertation demonstrate that information on heat network efficiency and flexibility, derived by modeling and analyzing highly dynamic user-level behavior captured in building-level data, can be obtained through appropriate modeling and analysis approaches. This cumulative dissertation presents valuable findings that advance highly dynamic user-level modeling and analysis in heat networks. They exemplify how user-level data can support efforts to reduce heat network temperatures and increase the capacity. The surplus knowledge gained by making highly dynamic user-level behavior accessible through modeling and analysis can enhance future heat network control. Integrating this type of information can enable practical steps to reduce safety margins without risking undersupply. Thus, this cumulative dissertation provides a fundamental knowledge base for practical implementation and scientific implications, serving as a starting point for future research to further enhance the efficiency and flexibility of heat networks.

## 5 References

- [1] ANTONOPOULOS, I./ROBU, V./COURAUD, B./KIRLI, D./NORBU, S./KIPRAKIS, A./FLYNN, D./ELIZONDO-GONZALEZ, S./WATTAM, S.: AI and ML approaches to energy demand-side response: A systematic review. *Renewable and Sustainable Energy Reviews*, 130, 109899, ISSN 1364-0321, <https://doi.org/10.1016/j.rser.2020.109899>, 2020.
- [2] BACHER, P./MADSEN, H./NIELSEN, H.A.: Online short-term heat load forecasting for single family houses. *IECON 2013 - 39th Annual Conference of the IEEE Industrial Electronics Society*, Vienna, Austria, 5741-5746, <https://doi.org/10.1109/IECON.2013.6700075>, 2013.
- [3] BACQUET, A./GALINDO FERNÁNDEZ, M./OGER, A./THEMESSL, N./FALLAHNEJAD, M./POPOVSKI, E./STEINBACH, J./BÜRGER, V./KÖHLER, B./BRAUNGARDT, S./BILLERBECK, A./BREITSCHOPF, B./WINKLER, J.: District heating and cooling in the European Union - Overview of markets and regulatory frameworks under the revised Renewable Energy Directive. EU publications, available online: <https://op.europa.eu/en/publication-detail/-/publication/4e28b0c8-eac1-11ec-a534-01aa75ed71a1/language-en>, 9:19 h, 27<sup>th</sup> October, 2022.
- [4] BAHLOWAN, H./FERRARO, N./GAMBAROTTA, A./LOSI, E./MANSERVIGI, L./MORINI, M./SALETTI, C./SPINA, P. R./VENTURINI, M.: Detection and identification of faults in a District Heating Network. *Energy Conversion and Management*, 266, 115837, <https://doi.org/10.1016/j.enconman.2022.115837>, 2022.
- [5] BALALI, Y./CHONG, A./BUSCH, A./O'KEEFE, S.: Energy modelling and control of building heating and cooling systems with data-driven and hybrid models—A review. *Renewable and Sustainable Energy Reviews*, 183, 113496, ISSN 1364-0321, <https://doi.org/10.1016/j.rser.2023.113496>, 2023.
- [6] BUFFA, S./FOULADFAR, M.H./FRANCHINI, G./LOZANO GABARRE, I./ANDRÉS CHICOTE, M.: Advanced Control and Fault Detection Strategies for District Heating and Cooling Systems—A Review. *Appl. Sci.*, 11, 455, <https://doi.org/10.3390/app11010455>, 2021.
- [7] BUJALSKI, M./MADEJSKI, P./FUZOWSKI, K.: Day-ahead heat load forecasting during the off-season in the district heating network using Generalized Additive model. *Energy and Buildings*, 278, 112630, ISSN 0378-7788, <https://doi.org/10.1016/j.enbuild.2022.112630>, 2023.
- [8] BUNDESMINISTERIUM FÜR JUSTIZ: Fernwärme- oder Fernkälte-Verbrauchserfassungs- und -Abrechnungsverordnung vom 28. September 2021 (BGBl. I S. 4591, 4831). Available online: <https://www.gesetze-im-internet.de/ffvav/BJNR459110021.html>, 09:43 h, 27<sup>th</sup> October 2022, Bundesamt für Justiz, 2021.
- [9] CALIKUS, E./NOWACZYK, S./SANT'ANNA, A./GADD, H./WERNER, S.: A data-driven approach for discovering heat load patterns in district heating. Elsevier, *Applied Energy*, 252, 113409, <https://doi.org/10.1016/j.apenergy.2019.113409>, 2019.
- [10] CHUJAI, P./KERDPRASOP, N./KERDPRASOP, K.: Time series analysis of household electric consumption with ARIMA and ARMA models. *Proceedings of the International Multi-Conference of Engineers and Computer Scientists*, I, 2013.

- [11] CHUNG, W. H./GU, Y. H./YOO, S. J.: District heater load forecasting based on machine learning and parallel CNN-LSTM attention. *Energy*, 246, 123350, ISSN 0360-5442, <https://doi.org/10.1016/j.energy.2022.123350>, 2022.
- [12] ĆIRIĆ, I./IGNJATOVIĆ, M./STOJILJKOVIĆ, M./STOJILJKOVIĆ, D./GOCIĆ, M./ĆIRIĆ, M.: Intelligent Heat Demand prediction for Advanced District Heat Plant Control. Zdravković, M./Konjović, Z./Trajanović, M. (Eds.), *ICIST 2020 Proceedings*, 198-201, 2020.
- [13] CUI, M.: District heating load prediction algorithm based on bidirectional long short-term memory network model. *Energy*, 254, Part A, 124283, ISSN 0360-5442, <https://doi.org/10.1016/j.energy.2022.124283>, 2022.
- [14] DAHL, M./BRUN, A./ANDRESEN, G. B.: Using ensemble weather predictions in district heating operation and load forecasting. *Applied Energy*, 193, 455-465, ISSN 0306-2619, <https://doi.org/10.1016/j.apenergy.2017.02.066>, 2017.
- [15] DANG, L.M./LEE, S./LI, Y./OH, C./NGUYEN, T. N./SONG, H./MOON, H.: Daily and seasonal heat usage patterns analysis in heat networks. *Sci. Rep.*, 12, 9165, <https://doi.org/10.1038/s41598-022-13030-6>, 2022.
- [16] DANG, L.M./SHIN, J./LI, Y./TIGHTIZ, L./NGUYEN, T. N./SONG, H./MOON, H.: Toward explainable heat load patterns prediction for district heating. *Sci. Rep.*, 13, 7434, <https://doi.org/10.1038/s41598-023-34146-3>, 2023.
- [17] DEUTSCHER WETTERDIENST: Index of/climate\_environment/CDC/observations\_germany/climate/hourly/. Available online: [https://opendata.dwd.de/climate\\_environment/CDC/observations\\_germany/climate/hourly/](https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/hourly/), 13:22 h, 14th July, 2024.
- [18] DEUTSCHER WETTERDIENST: Wetter und Klima – Deutscher Wetterdienst – Qualität Wettervorhersage. [https://www.dwd.de/DE/wetter/schon\\_gewusst/qualitaetvorhersage/qualitaetvorhersage\\_node.html](https://www.dwd.de/DE/wetter/schon_gewusst/qualitaetvorhersage/qualitaetvorhersage_node.html), accessed 27th March, 2024.
- [19] DEUTSCHER WETTERDIENST: Wettervorhersage - Wetter\_modell\_Prognoseguete.pdf. [https://www.dwd.de/DE/fachnutzer/landwirtschaft/dokumentationen/agrowetter/Wetter\\_Modell\\_Prognoseguete.pdf?\\_\\_blob=publicationFile&v=2](https://www.dwd.de/DE/fachnutzer/landwirtschaft/dokumentationen/agrowetter/Wetter_Modell_Prognoseguete.pdf?__blob=publicationFile&v=2), accessed 27th March, 2024.
- [20] FINKENRATH, M./FABER, T./BEHRENS, F./LEIPRECHT, S.: Holistic modelling and optimisation of thermal load forecasting, heat generation and plant dispatch for a district heating network. *Energy*, 250, 123666, ISSN 0360-5442, <https://doi.org/10.1016/j.energy.2022.123666>, 2022.
- [21] GADD H./WERNER S.: Fault detection in district heating substations. *Applied Energy* 157, 51-59, <https://doi.org/10.1016/j.apenergy.2015.07.061>, 2015.
- [22] GADD, H./WERNER, S.: Achieving low return temperatures from district heating substations. *Applied Energy*, 136, 59-67, 2014.
- [23] GERMAN FEDERAL STATISTICAL OFFICE (DESTATIS): Erneuerbare Energien in ausgewählten Sektoren in der EU-27 – Anteil am Bruttoenergieverbrauch in %. 2026, available online: [https://www.destatis.de/Europa/DE/Thema/GreenDeal/\\_Grafik/03-erneuerbare-energie-sektoren.html](https://www.destatis.de/Europa/DE/Thema/GreenDeal/_Grafik/03-erneuerbare-energie-sektoren.html), last downloaded 17:33h, 3rd February, 2026.
- [24] GEYSEN, D./DE SOMER, O./JOHANSSON, C./BRAGE, J./VANHOUDT, D.: Operational thermal load forecasting in district heating networks using ML and expert advice. *Energy*

- and Buildings, 162, 144-153, ISSN 0378-7788, <https://doi.org/10.1016/j.enbuild.2017.12.042>, 2018.
- [25] GIANNIOU, P./LIU, X./HELLER, A./NIELSEN, P. S./RODE, C.: Clustering-based analysis for residential district heating data. *Energy Conversion and Management*, 165, 840–850, 2018.
- [26] GOLLA, A./GEIS, J./LOY, T./STAUDT, P./WEINHARDT, C.: An operational strategy for district heating networks: application of data-driven heat load forecasts. 9th DACH+ Conference on Energy Informatics, Sierre, Switzerland, 29–30 October 2020, *Energy Informatics*, 3, Suppl. 1, 1-22, <https://doi.org/10.1186/s42162-020-00125-5>, 2020.
- [27] GONG, M./ZHAO, Y./SUN, J./HAN, C./SUN, G./YAN, B.: Load forecasting of district heating network based on Informer. *Energy*, 253, 124179, ISSN 0360-5442, <https://doi.org/10.1016/j.energy.2022.124179>, 2022.
- [28] GREFF, K./SRIVASTAVA, R. K./KOUTNÍK, J./STEUNEBRINK, B. R./SCHMIDHUBER, J.: LSTM: A Search Space Odyssey. *IEEE Transactions on Neural Networks and Learning Systems*, 28, no. 10, 2222-2232, 2017, <https://doi.org/10.1109/TNNLS.2016.2582924>, 2017.
- [29] GROß, A./LENDERS, A./SCHWENKER, F./BRAUN, D. A./FISCHER, D.: Comparison of short-term electrical load forecasting methods for different building types. *Energy Inform.*, 4, Suppl 3, 13, <https://doi.org/10.1186/s42162-021-00172-6>, 2021.
- [30] HOCHREITER, S./ADLER, T.: LSTM and Recurrent Neural Nets - Lecture Notes. Institute for Machine Learning, Johannes Kepler University Linz, Winter Semester, 2021.
- [31] HOU, H./LIU, C./WANG, Q./WU, X./TANG, J./SHI, Y./XIE, C.: Review of load forecasting based on AI methodologies, models, and challenges. *Electric Power Systems Research*, 210, 108067, ISSN 0378-7796, <https://doi.org/10.1016/j.epsr.2022.108067>, 2022.
- [32] IDOWU, S./SAGUNA, S./ÅHLUND, C./SCHELÉN, O. Forecasting heat load for smart district heating networks: A machine learning approach. 2014 IEEE International Conference on Smart Grid Communications (SmartGridComm), Venice, Italy, 554-559, <http://doi.org/10.1109/SmartGridComm.2014.7007705>, 2014.
- [33] IDOWU, S./SAGUNA, S./ÅHLUND, C./SCHELÉN, O.: Applied ML: Forecasting heat load in district heating network. *Energy and Buildings*, 133, 478-488, <https://doi.org/10.1016/j.enbuild.2016.09.068>, 2016.
- [34] JOHANSSON, C./BERGKVIST, M./GEYSEN, D./DE SOMER, O./LAVESSON, N./VANHOUDT, D.: Operational Demand Forecasting In District Heating Networks Using Ensembles Of Online Machine Learning Algorithms. *Energy Procedia*, 116, 208-216, <https://doi.org/10.1016/j.egypro.2017.05.068>, 2017.
- [35] KEMPER, N./HEIDER, M./PIETRUSCHKA, D./HÄHNER, J.: Forecasting of residential unit's heat demands: a comparison of machine learning techniques in a real-world case study. *Energy Systems*, 16, 281 – 315, <https://doi.org/10.1007/s12667-023-00579-y>, 2025.
- [36] KEYAN, L./DONGLI, J./ZHONGJIAN, K./LUO, L.: Anomaly Detection Method of Distribution Network Line Loss Based on Hybrid Clustering and LSTM. *J. Electr. Eng. Technol.*, 17, 1131–1141, <https://doi.org/10.1007/s42835-021-00958-4>, 2022.

- [37] KIM, J./KIM, S./CHOI, S.: Learning to Warm-Start Bayesian Hyperparameter Optimization. arXiv:1710.06219, Version 3, <https://doi.org/10.48550/arXiv.1710.06219>, 2018.
- [38] KUDELA, L./CHÝLEK, R./POSPÍŠIL, J.: Efficient Integration of ML into District Heating Predictive Models. *Energies*, 13, 6381, <https://doi.org/10.3390/en13236381>, 2020.
- [39] KUREK, T./BIELECKI, A./ŚWIRSKI, K./WOJDAN, K./GUZEK, M./BIAŁEK, J./BRZOZOWSKI, R./SERAFIN, R.: Heat demand forecasting algorithm for a Warsaw district heating network. *Energy*, 217, 119347, ISSN 0360-5442, <https://doi.org/10.1016/j.energy.2020.119347>, 2021.
- [40] LAAKKONEN, L./KORPELA, T./KAIVOSOJA, J./VILKKO, M./MAJANNE, Y./NURMORANTA, M.: Predictive Supply Temperature Optimization of District Heating Networks Using Delay Distributions. *Energy Procedia*, 116, 297-309, <https://doi.org/10.1016/j.egypro.2017.05.076>, 2017.
- [41] LEE, D./YOON, S.M./LEE, J./KIM, K./SONG, S.H.: Applying Deep Learning to the Heat Infeed Planning Problem in a District Heating network. *Energies*, 13, 6641, <https://doi.org/10.3390/en13246641>, 2020.
- [42] LEIPRECHT, S./BEHRENS, F./FABER, T./FINKENRATH, M.: A comprehensive thermal load forecasting analysis based on ML algorithms. *Energy Reports*, 7, Supplement 4, 319-326, ISSN 2352-4847, <https://doi.org/10.1016/j.egyr.2021.08.140>, 2021.
- [43] LIN, T./PAN, Y./XUE, G./SONG, J./QI, C.: A Novel Hybrid Spatial-Temporal Attention-LSTM Model for Heat Load Prediction. *IEEE Access*, 8, 159182-159195, <https://doi.org/10.1109/ACCESS.2020.3017516>, 2020.
- [44] LUMBRERAS, M./GARAY-MARTINEZ, R./ARREGI, B./MARTIN-ESCUDERO, K./DIARCE, G./RAUD, M./HAGU, I.: Data driven model for heat load prediction in Buildings connected to District Heating by using smart heat meters. *Energy*, 239, Part D, 122318, <https://doi.org/10.1016/j.energy.2021.122318>, 2022.
- [45] LUND, H./WERNER, S./WILTSHIRE, R./SVENDSEN, S./THORSEN, J. E./HVELPLUND, F./MATHIESEN, B. V.: 4th Generation District Heating (4GDH): Integrating smart thermal grids into future sustainable energy systems. *Energy*, 68, 1-11, <https://doi.org/10.1016/j.energy.2014.02.089>, 2014.
- [46] LV, R./YUAN, Z./LEI, B./ZHENG, J./LUO, X.: Building thermal load prediction using deep learning method considering time-shifting correlation in feature variables. *Journal of Building Engineering*, 61, 105316, <https://doi.org/10.1016/j.jobbe.2022.105316>, 2022.
- [47] MAHAJAN, T./SINGH, G./BRUNS, G./BRUNS, G./MAHAJAN, T./SINGH, G.: An experimental assessment of treatments for cyclical data. In *Proceedings of the 2021 Computer Science Conference for CSU Undergraduates*, 6, 22, <http://hdl.handle.net/20.500.12680/th83m446n>, 2021.
- [48] MÅNSSON, S./JOHANSSON KALLIONIEMI, P.-O./SERNHED, K./THERN, M.: A machine learning approach to fault detection in district heating substations. *Energy Procedia*, 149, 226-235, <https://doi.org/10.1016/j.egypro.2018.08.187>, 2018.

- [49] MÄNSSON, S./JOHANSSON KALLIONIEMI, P.-O./THERN, M./VAN OEVELEN, T./SERNHED, K.: Faults in district heating customer installations and ways to approach them: Experiences from Swedish utilities. *Energy*, 180, 163-174, <https://doi.org/10.1016/j.energy.2019.04.220>, 2019.
- [50] MATHWORKS: Partition time series data for cross-validation – MATLAB - MathWorks Deutschland. Available online: <https://de.mathworks.com/help/stats/tspartition.html>, 11:06 h, 30th September, 2024.
- [51] MBIYDZENYUY, G./NOWACZYK, S./KNUTSSON, H./VANHOUDT, D./BRAGE, J./CALIKUS, E.: Opportunities for Machine Learning in District Heating. *Appl. Sci.*, 11, 6112, <https://doi.org/10.3390/app11136112>, 2021.
- [52] NAIK, K./IANAKIEV, A.: Heat demand prediction: A real-life data model vs simulated data model comparison. *Energy Reports*, 7, Supplement 4, 380-388, ISSN 2352-4847, <https://doi.org/10.1016/j.egyr.2021.08.093>, 2021.
- [53] NEUMAYER, M./STECHEER, D./GRIMM, S./MAIER, A./BÜCKER, D./SCHMIDT, J.: Fault and anomaly detection in district heating substations: A survey on methodology and data sets. *Energy*, 276, 127569, <https://doi.org/10.1016/j.energy.2023.127569>, 2023.
- [54] NTAKOLIA, C./ANAGNOSTIS, A./MOUSTAKIDIS, S./KARCANIAS, N.: ML applied on the district heating and cooling sector: a review. *Energy Syst.*, 13, 1–30, <https://doi.org/10.1007/s12667-020-00405-9>, 2022.
- [55] POTOČNIK, P./ŠKERL, P./GOVEKAR, E.: Machine-learning-based multi-step heat demand forecasting in a district heating network. *Energy and Buildings*, 233, 110673 <https://doi.org/10.1016/j.enbuild.2020.110673>, 2021.
- [56] POTOČNIK, P./VIDRIH, B./KITANOVSKI, A./GOVEKAR, E.: Neural network, ARX, and extreme learning machine models for the short-term prediction of temperature in buildings. University of Ljubljana, Faculty of Mechanical Engineering, *Build. Simul.*, 12, 1077–1093, <https://doi.org/10.1007/s12273-019-0548-y>, 2019.
- [57] SAKKAS, N.P./ABANG, R.: Thermal load prediction of communal district heating networks by applying data-driven ML methods. *Energy Reports*, 8, 1883-1895, ISSN 2352-4847, <https://doi.org/10.1016/j.egyr.2021.12.082>, 2022.
- [58] SALA, J./LI, R./CHRISTENSEN, M.H.: Clustering and classification of energy meter data: A comparison analysis of data from individual homes and the aggregated data from multiple homes. *Build. Simul.*, 14, 103–117, <https://doi.org/10.1007/s12273-019-0587-4>, 2021.
- [59] SALOUX, E./CANDANEDO, J.A.: Forecasting District Heating Demand using ML Algorithms. *Energy Procedia*, 149, 59-68, ISSN 1876-6102, <https://doi.org/10.1016/j.egypro.2018.08.169>, 2018.
- [60] SANDIN, F./GUSTAFSSON, J./DELSING, J./EKLUND, R.: Basic methods for automated fault detection and energy data validation in existing district heating networks. 13th International Symposium on District Heating and Cooling 3rd of September - 4th of September Copenhagen Denmark, <https://urn.kb.se/resolve?urn=urn:nbn:se:ltu:diva-40519>, 2012.
- [61] SHAMSHIRBAND, S./PETKOVIĆ, D./ENAYATIFAR, R./ABDULLAH, A. H./MARKOVIĆ, D./LEE, M./AHMAD, R.: Heat load prediction in district heating networks with adaptive

- neuro-fuzzy method. *Renewable and Sustainable Energy Reviews*, 48, 760-767, ISSN 1364-0321, <https://doi.org/10.1016/j.rser.2015.04.020>, 2015.
- [62] SIENCEDIRECT: Science, health and medical journals, full text articles and books. Available online: <https://www.sciencedirect.com/>, 19th and 20th September, 2022.
- [63] SINGH, D./BIRMOHAN SINGH, B.: Investigating the impact of data normalization on classification performance. *Applied Soft Computing*, 97, Part B, 105524, ISSN 1568-4946, <https://doi.org/10.1016/j.asoc.2019.105524>, 2020.
- [64] SONG, J./ZHANG, L./XUE, G./MA, Y.P./GAO, S./JIANG, Q.L.: Predicting hourly heating load in a district heating network based on a hybrid CNN-LSTM model. *Energy and Buildings*, 243, 110998, ISSN 0378-7788, <https://doi.org/10.1016/j.enbuild.2021.110998>, 2021.
- [65] SURYANARAYANA, G./LAGO, J./GEYSEN, D./ALEKSIEJUK, P./JOHANSSON, C.: Thermal load forecasting in district heating networks using deep learning and advanced feature selection methods. *Energy*, 157, 141-149, <https://doi.org/10.1016/j.energy.2018.05.111>, 2018.
- [66] VANTING, N.B./MA, Z./JØRGENSEN, B.N.: Evaluation of neural networks for residential load forecasting and the impact of systematic feature identification. *Energy Inform.*, 5, Suppl. 4, 63, <https://doi.org/10.1186/s42162-022-00224-5>, 2022.
- [67] WANG, C./YUAN, J./HUANG, K./ZHANG, J./ZHENG, L./ZHOU, Z./ZHANG, Y.: Research on thermal load prediction of district heating station based on transfer learning. *Energy*, 239, Part E, 122309, ISSN 0360-5442, <https://doi.org/10.1016/j.energy.2021.122309>, 2022.
- [68] WEI, Z./ZHANG, T./YUE, B./DING, Y./XIAO, R./WANG, R./ZHAI, X.: Prediction of residential district heating load based on ML: A case study. *Energy*, 231, 120950, ISSN 0360-5442, <https://doi.org/10.1016/j.energy.2021.120950>, 2021.
- [69] WORDEN, K./DULIEU-BARTON, J. M.: An overview of intelligent fault detection in systems and structures. *Struct. Health Monit.*, 3(1), 85–98, <http://dx.doi.org/10.1177/1475921704041866>, 2004.
- [70] XUE, G./QI, C./LI, H./KONG, X./SONG, J.: Heating load prediction based on attention long short term memory: A case study of Xingtai. *Energy*, 203, 117846, ISSN 0360-5442, <https://doi.org/10.1016/j.energy.2020.117846>, 2020.
- [71] XUE, P./JIANG, Y./ZHOU, Z./CHEN, X./FANG, X./LIU, J.: Multi-step ahead forecasting of heat load in district heating networks using ML algorithms. *Energy*, 188, 116085, ISSN 0360-5442, <https://doi.org/10.1016/j.energy.2019.116085>, 2019.
- [72] YAO, F./ZHOU, W./AL GHAMDI, M./SONG, Y./ZHAO, W.: An integrated D-CNN-LSTM approach for short-term heat demand prediction in district heating networks. *Energy Reports*, 8, Supplement 13, 98-107, ISSN 2352-4847, <https://doi.org/10.1016/j.egyr.2022.08.087>, 2022.
- [73] YU, Z./FUNG, B./HAGHIGHAT, F.: Extracting knowledge from building-related data — A data mining framework. *Build. Simul.*, 6, 207–222, <https://doi.org/10.1007/s12273-013-0117-8>, 2013.

- [74] ZDRAVKOVIĆ, M./ĆIRIĆ, I./IGNJATOVIĆ, M.: Explainable heat demand forecasting for the novel control strategies of district heating networks. *Annual Reviews in Control*, 53, 405-413, ISSN 1367-5788, <https://doi.org/10.1016/j.arcontrol.2022.03.009>, 2022.
- [75] ZDRAVKOVIĆ, M./ĆIRIĆ, I./IGNJATOVIĆ, M.: Towards explainable AI-assisted operations in District Heating networks. *IFAC-PapersOnLine*, 54, Issue 1, 390-395, ISSN 2405-8963, <https://doi.org/10.1016/j.ifacol.2021.08.044>, 2021.
- [76] 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*, 285, 116452, ISSN 0306-2619, <https://doi.org/10.1016/j.apenergy.2021.116452>, 2021.
- [77] ZHAO, A./MI, L./XUE, X./XI, J./JIAO, Y.: Heating load prediction of residential district using hybrid model based on CNN. *Energy and Buildings*, 266, 112122, ISSN 0378-7788, <https://doi.org/10.1016/j.enbuild.2022.112122>, 2022.
- [78] ZHAO, B./JIN, Y./LI, W./ZHENG, H.: Analysis on the Technical Situation and Applied Difficulties of District Heating Load Forecasting. *Therm. Eng.*, 69, 464-472, <https://doi.org/10.1134/S0040601522060088>, 2022.
- [79] ZHAO, Y./LI, T./ZHANG, X./ZHANG, C.: Artificial intelligence-based fault detection and diagnosis methods for building energy systems: Advantages challenges and the future. *Renewable and Sustainable Energy Reviews*, 109, 85-101, <https://doi.org/10.1016/j.rser.2019.04.021>, 2019.
- [80] ZHONG, W./FENG, E./LIN, X./XIE, J.: Research on data-driven operation control of secondary loop of district heating network. *Energy*, 239, Part B, 122061, <https://doi.org/10.1016/j.energy.2021.122061>, 2022.