Final report of the doctoral scholarship
of the Society for Energy and Climate Protection
(Gesellschaft für Energie- und Klimaschutz)
Schleswig-Holstein GmbH

Potentials of optimization and flexibilization in pipe-bound heat systems

---

by
M.Eng. Samanta Weber

Supervision of the doctoral thesis:
Prof. Dr. mont. Dr. rer. nat. Michael Fischlschweiger
University of Technology Clausthal
and
Prof. Dr.-Ing. Dirk Volta
University of Applied Sciences Flensburg

Funding and support:
Stadtwerke Flensburg GmbH
and
Gesellschaft für Energie- und Klimaschutz
Schleswig-Holstein GmbH

TU Clausthal

Hochschule Flensburg
University of Applied Sciences

stadtwerke flensburg

EKSH
Gesellschaft für
Energie und Klimaschutz
Schleswig-Holstein

Table of contents

List of Figures

List of Abbreviations

DL DEEP LEARNING
ML MACHINE LEARNING

Abstract

Recent climate events have accelerated the efforts to reach the climate-neutral operation of energy systems, including a critical focus on heat networks. For the climate-neutral operation of existing heat networks, one key requirement is increased efficiency and flexibility. Machine Learning (ML) has shown promising results in enhancing the efficiency and flexibility of thermal energy systems. While current research mainly focuses on heat network input parameters, the thesis related to this documentation proposes using user-level data to approach this objective further. It investigates how substation data from a heat network in northern Germany can be integrated with ML models to achieve efficiency and flexibility gains.

The research follows three phases linked to associated research thematics. The thesis related to this documentation aims to develop: a) models for user-level parameters, i.e., the volume flow, the supply temperature, and the return temperature for application in enhanced heat network control systems, and b) a fault management method for user-level substations.

Research Thematic I identifies key features for ML modeling of the named user-level parameters, with the objective of accelerating ML application in heat networks and providing knowledge for Research Thematic III.

Research Thematic II establishes an automated method for detecting faulty substations and quantifying their (negative) impact on the system. The sub-objective is twofold: to enhance network efficiency and to enable the exclusion of fault-inflicted data for modeling within Research Thematic III.

Research Thematic III investigates the model-building process for the highly dynamic user-level parameters, with a particular focus on data pre-processing, optimization of the model hyperparameters, and impact evaluation of domain-specific model settings. The models are tested on substation data. Excluding substations with fault-inflicted or otherwise anomalous patterns improves modeling success. The models could reach error metric values comparable to those in the existing literature despite the high dynamics of the user-level data.

The findings show that substations with repetitive, reliable patterns can be successfully modeled, while those with random behavior or faults still induce challenges. The fault management system effectively detects affected substations and quantifies the negative influence of the fault on the thermal energy system, enhancing data labeling and network efficiency. Further research could refine the models for advancing heat network control and efficiency. The ML's potential is demonstrated on the user level in heat networks, providing new approaches for efficiency and flexibility gains, including fault management and model building for enhanced heat network control.

1 Introduction

Heat networks have become a key focus, as they are widely regarded as the most efficient heat supply solution in areas with high demand density [9, 19, 21]. To meet future requirements, existing heat networks should achieve lower operational temperatures and greater flexibility, as addressed in the concept of the 4th generation of district heating [24]. ML has emerged as a promising tool for simulating and enhancing efficiency and flexibility in thermal energy systems [31]. The application of data-driven methods to heat networks is relatively recent and presents new opportunities, accompanied by challenges in heat network modeling [14].

Data-driven models allow for automated model building [29], facilitating adaptation to different heat networks. Nonetheless, limited data availability remains a significant inhibitor to data-driven modeling [3]. Current research often focuses on the available data, and, thus, on modeling the heat systems' inputs, particularly in heat load prediction (e.g., Ref. [7, 9, 10, 11, 12, 14, 17, 21, 23, 25, 33, 37, 40, 42]. In comparison, user-level data has only recently become accessible. In Germany, for example, remote data collection has been required for newly installed systems since 2021, with retrofits mandatory by 2027 [6]. This study uses data from a model region in Tarp, Schleswig-Holstein, Germany. The data include recordings from the infeed facility and user-level substations, such as volume flow, supply temperature, and return temperature, in approximately hourly resolution.

Given the limited research on using ML methods with user-level heat network data [29], the doctoral thesis associated with this documentation aims to contribute to filling that gap. Two major research directions have been identified regarding the potential for efficiency and flexibility gains using ML with user-level data, which are outlined below.

1.1 Modeling of user-level parameters for enhanced heat network control

Current research on ML modeling (including Deep Learning (DL)) in heat networks primarily inquires about heat load prediction [29]. However, practical boundary conditions, such as the risk of undersupply, need to be comprehensively addressed [29], which limits the applicability. One suggested solution is to close the information gap between operators and users by implementing a bottom-up control approach for heat networks [29, 45]. More precise predictions for substation operational parameters could reduce safety margins [29], minimizing losses without taking the risk of undersupply.

Nevertheless, the unique nature of experimental data requires specific information for optimal model development, such as the selection of inputs [10, 11, 25, 33, 40, 42], data pre-processing methods [11, 12, 15, 16, 26, 32, 35, 43], modeling architecture [31], hyperparameter tuning [22, 36], and further domain-specific settings. This knowledge, particularly at the user level in heat networks, needs improvement [3]. Therefore, generating information for predicting user-level heat network parameters is essential to enhance the effectiveness of ML applications, aiming to enhance efficiency and flexibility in heating systems.

1.2 Effective fault management on the substation level

A second area of focus for improving heat system efficiency at the user level is fault management [30], especially for individual substations [29]. Faults often require either an increase in volume flow or a rise in supply temperature to maintain user comfort [27], decreasing the overall system efficiency. Effective fault management is essential, as it can account for up to half of the temperature reduction needed to achieve the efficiency standards of 4th generation district heating [13]. However, operators have limited resources for managing faults [4], making it essential to reduce manual processing steps in fault detection [4], e.g., through automation. Additionally, assigning a quantifiable measure of a fault's impact [28] allows for prioritizing fault elimination measures. Fault management methods are requested to rely on accessible data, such as primary-side user measurements [34], and be reliable, cost-effective, and adaptable to various heat network scales [44]. There is currently a lack of methods that meet these overarching objectives. Therefore, the thesis related to this documentation aims to address this gap and contribute to improving heating system efficiency and flexibility by achieving corresponding implementation-level sub-objectives.

Fault management at the user level also affects data quality, which is critical for ML model development. Sufficient clean data are essential for accurate ML modeling [3]. Yet, faults in substations can lead to recordings that do not reflect the expected behavior in general and efficiency in specific, resulting in misleading ML models. Identifying and excluding such faulty datasets is necessary but also reduces the amount of usable data for modeling.

1.3 Resulting research question and research steps

To fill the identified research gaps, the thesis related to this documentation investigates the research question preliminarily formulated as follows: How can ML architectures and available user-level heat network data be combined with a high degree of automation to enhance a heating system's efficiency and flexibility beyond the state-of-the-art?

Hence, this study contributes three novel research thematics on data-driven efficiency and flexibility enhancements in heat networks. Further, the aim already partially achieved is to publish the findings in international peer-reviewed scientific journals. The three research thematics are summarized below.

1. Feature Engineering and Selection: The first step in ML modeling is to identify the appropriate input features for predicting the desired outputs. The process of transforming raw data into processible datasets for ML is known as "feature engineering," while "feature selection" means choosing a subset of relevant and desirably non-redundant features. These steps are critical for successful modeling [33] but are time-consuming and investigated to a limited extent for heat networks [10, 25], particularly at the user level. The thesis associated with this documentation aims to establish a suitable set of features for the proposed modeling targets, accelerating the model-building process and enhancing the applicability of ML in heat networks.

2. Fault Management at the Substation Level: Fault management is crucial, as faults can increase heat losses and, thus, reduce the overall system efficiency [29] and user comfort [2, 41]. Given that the number of substations exhibiting faults can be significant (e.g., 74 % [13]), monitoring and coordinating measures for fault elimination are essential. Existing methods offer fault management solutions but often do not adequately account for practical constraints faced by operators [4, 34, 44]. This study introduces a three-step fault management process designed to integrate practical requirements, e.g., using only primary-side data from user substations [34], ensuring high automation [4] by limiting the manual processing steps, and providing a quantifiable impact assessment of detected faults [28]. This method is also designed for adaptation to other heat networks [44] by stating the required adjustment steps, making it a practical and potentially transferable solution.

3. Data pre-processing, hyperparameter tuning, and domain-specific model settings: Before ML models can be trained, data must be pre-processed, which includes tasks such as time-stamp equalization, outlier removal, missing data filling [8], encoding non-numerical data (e.g., human behavior [10, 11, 17, 20, 33, 37]), as well as scaling and data transformation [35, 43]. These steps

significantly impact modeling success [35], but research in this area is limited for heat network parameters [43], especially on the user level with the inherent high dynamics. Thus, this study explores data scaling methods and the transformation of cyclic continuous features with sine and cosine [12, 26] to assess their impact on training and model performance. Additionally, ML models require the manual setting of hyperparameters, which are crucial for optimal performance [22] but remain researched to a limited extent for heat networks [43]. Methods such as Bayesian optimization allow the transfer of information on suitable hyperparameter ranges from comparable modeling tasks [18]. The literature currently explores specific hyperparameters (e.g., Ref. [22, 36]), but identifying suitable hyperparameter ranges for optimization at the user level in heat networks remains explored to a limited extent. Additionally, domain-specific factors influencing modeling performance—such as seasonality [1, 5, 20] and data quantity [5] —present challenges that have not yet been fully addressed in existing research. To fill this gap, the thesis focuses on data pre-processing and hyperparameter tuning, assessing the impact of various model settings, including domain-specific influences for user-level parameters as modeling targets. This aims to provide information to enhance automation in user-level modeling for heat networks.

2 Research and Results

Until recently, limited data availability has prevented comprehensive research in applying ML to user-level data in heat networks. Nevertheless, existing studies (e.g., Ref. [20, 29, 38]) have demonstrated the effectiveness of ML in enhancing the efficiency and flexibility of heat networks, indicating that user-level data could be crucial for further optimizing thermal energy systems.

Based on this, the objective guiding the thesis associated with this documentation is preliminarily formulated as follows: With ML algorithms, information essential for exploiting the existing potential for optimization in heat networks can be accessed by relying on the available experimental data. Thus, the now additionally available user data can reveal further potential for increasing the efficiency and flexibility of thermal energy systems by applying comparable ML approaches.

Therefore, the study develops a methodological approach to collect the necessary information for using ML with user-level thermal energy system data. The measured user-level data, including volume flow, supply temperature, and return temperature, can provide critical information on demand and supply safety, bridging the information gap between the heat network operator and the user. These parameters are, thus, proposed as new prediction targets for enhancing heat network control. Additionally, these data contain insights into substation efficiency, enabling the use in fault management to improve both network efficiency and flexibility as well as data quality. The essential steps include an iterative process to find the most successful combination of data and modeling architecture for applying ML algorithms in this context, as given in Figure 1.

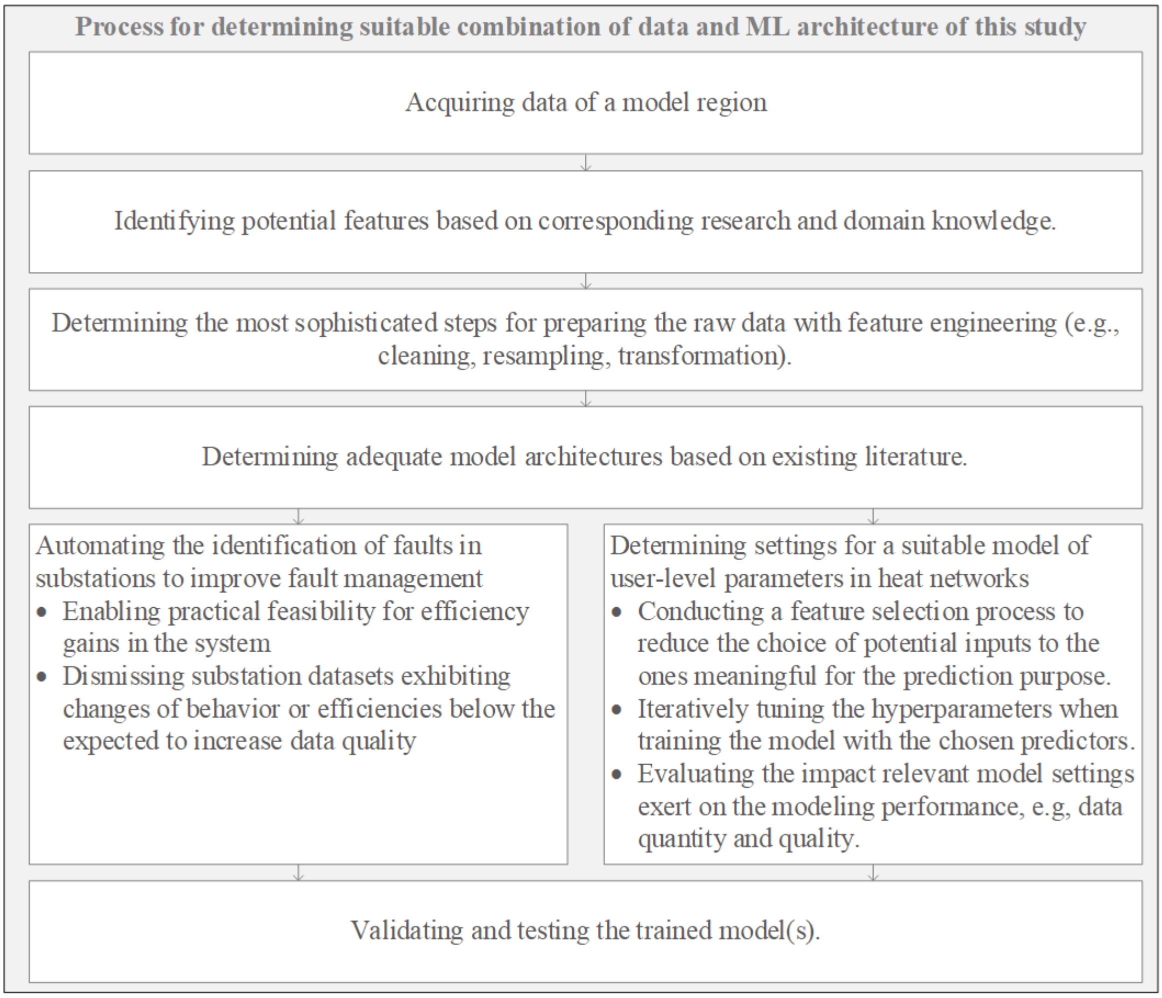

Figure 1: Steps for determining a suitable combination of data and modeling architecture

A three-phase research design was developed to gather the necessary information. Each phase is associated with a research thematic addressing a specific aspect of the research question. For Research Thematic I, feature engineering and selection were reviewed to assess existing knowledge on modeling user data in this context. Features used in previous studies for heat network modeling were identified. Research Thematic I proposes a methodological approach for feature engineering as well as investigation and selection, offering a set of relevant, non-redundant features for modeling user-level data in the model region. For Research Thematic II, the suitability of the accessible user-level data for fault management was tested. Research Thematic II presents a three-step approach for fault management, focusing on practical feasibility and enabling automated fault detection and impact evaluation in substations. This method allows for an automated exclusion of substation datasets unsuitable for user-level modeling, aligning with the study's objectives. For Research Thematic III, the research findings were combined to build ML models for the newly proposed prediction targets. Research Thematic III discusses data pre-processing, hyperparameter tuning, and the effects of domain-specific hyperparameter settings on modeling success. It then evaluates the validity of the results based on models for substations with adequate data collection rates.

In terms of industrial knowledge transfer, the study aims to enhance understanding of the model region’s heat network and to contribute new knowledge on thermal energy systems by introducing new prediction targets. The ML-based approach increases the potential to generalize and adapt the findings to other heat networks. By defining the limitations and potentials of the methods, the study related to this documentation validates the suitability of these approaches for achieving efficiency and flexibility gains for one particular heat network. The following sections provide an overview of the three research thematics, highlighting results, scientific contributions, and the relation to the addressed research questions.

2.1 Research Thematic I: Feature Engineering and Selection

2.1.1 Summary of Research Thematic I

For Research Thematic I, a review of existing literature was conducted to identify predictors used in ML modeling within heat networks. By combining domain knowledge and a comprehensive literature review, factors influencing the targeted user-level parameters were determined, as illustrated in Figure 2.

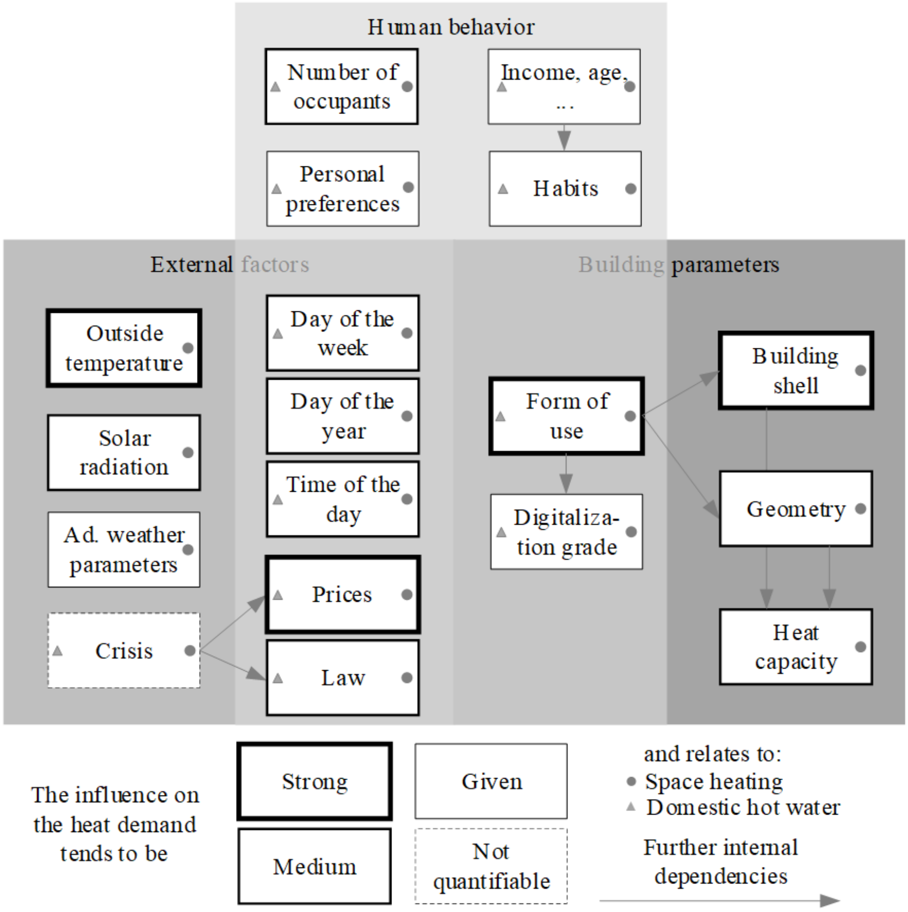

Figure 2: Influencing factors on the heat demand determined in literature for Research Thematic I

The assumption is that these influencing factors, mainly originating from heat load prediction, can be applied to user-level parameters as well. For this, raw data on the influencing factors must be available and converted into a numerical time series. As a result, a subset of potential features is extracted from these influencing factors. Following this, a methodology is developed to systematically identify features that are both relevant to the prediction targets and minimally redundant. This process establishes a sophisticated workflow and methodically implements the necessary processing steps. Overall, Research Thematic I aims to identify a set of appropriate features, potentially unique to the model region, for modeling the newly proposed prediction targets.

2.1.2 Publications related to Research Thematic I

- WEBER, S./FISCHLSCHWEIGER, M./VOLTA, D./GEISLER, J.: Feature engineering for machine learning to predict heat networks on the end-user level. 10th International Conference on Smart Energy Systems, Aalborg, 10th – 11th September 2024.
- Submitted: WEBER, S./FISCHLSCHWEIGER, M./VOLTA, D./GEISLER, J.: Statistical and Machine Learning based Feature Engineering and Selection for a heat network in the context of new targets of prediction for a model region. Article, Scientific Reports.

2.2 Research Thematic II: Fault Management

2.2.1 Summary of Research Thematic II

Research Thematic II explicitly addresses fault management in user substations as another critical aspect of improving the efficiency and flexibility of heat networks.

First, the literature is thoroughly reviewed to identify effective approaches for utilizing primary-side user data for fault detection. This review reveals a need to enhance the practical feasibility of existing fault management methods. Clustering is identified as a common methodological approach. Next, careful consideration is given to selecting the input data. The system temperature plays a critical role in determining the efficiency of the system, leading to the decision to reduce the required information for fault detection to just the temperature measurements. To improve practical feasibility and reduce dimensionality, the aim is to limit time dependency (information gained from the sequential order of the time series data), as seen in related research. Third, a systematic workflow and appropriate methods for implementation are selected. Finally, the resulting three-step method is tested and validated using data from the model region to assess the hypothesis of this research.

Fault detection rests on measurements of supply and return temperatures. The volume flow information is further integrated to quantify the faults' impact on the heating system. Figure 3 gives an exemplary visualization of the fault detection and impact quantification results (arbitrary, for data privacy reasons).

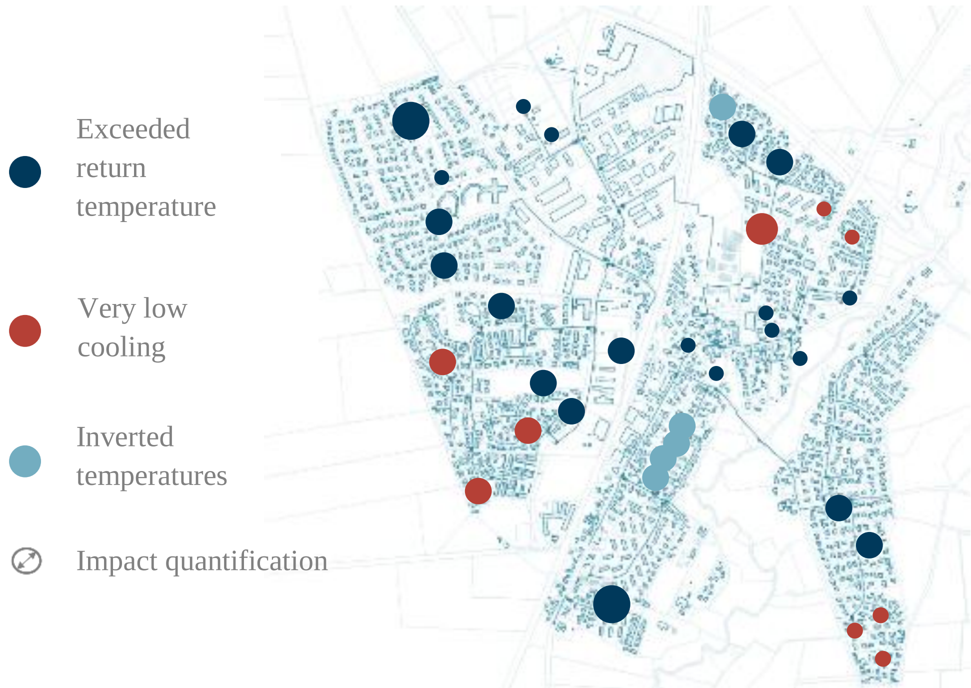

Figure 3: Exemplification of the fault detection and impact quantification results of Research Thematic II from [39]

Overall, Research Thematic II aimed to develop a fault management approach that would enhance system efficiency and flexibility while also improving data quality for Research Thematic III.

2.2.2 Publications related to Research Thematic II

- WEBER, S./FISCHLSCHWEIGER, M./VOLTA, D./RIECK-BLANKENBURG, U.: Identifikation und Einflussbewertung von Fehlern in Übergabestationen auf Endnutzerebene in Wärmenetzen. 2. Norddeutsche Wärmefachkonferenz, Hamburg, 19th – 20th September 2024.
- WEBER, S./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. Scientific Reports, 14, 32166, https://doi.org/10.1038/s41598-024-82103-5, 2024.

2.3 Research Thematic III: Data pre-preparation, Hyperparameters, Prediction

2.3.1 Summary of Research Thematic III

Research Thematic III focuses on acquiring the additional knowledge necessary to build a data-driven model for the newly proposed highly dynamic user-level parameters as prediction targets, with the overarching aim of enhancing the efficiency and flexibility of heat networks.

Research Thematic III rests on the features found within Research Thematic I. To identify an appropriate modeling architecture, a comprehensive literature review is conducted on settings for ML modeling in heat networks. Figure 4 presents the LSTM-network layers for the chosen basic structure.

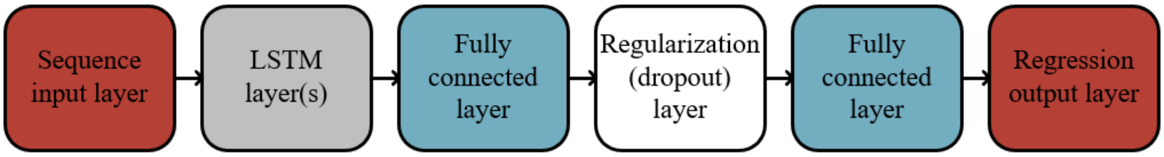

Figure 4: LSTM-network layers for basic model architecture with the number of LSTM layers part of hyperparameter optimization as applied within Research Thematic III

The literature review also outlines adequate data pre-processing methods such as scaling and transformation, a suitable approach to hyperparameter tuning—including relevant hyperparameters and their respective ranges for testing—and domain-specific hyperparameter settings to be examined individually. These settings include factors like data quantity and seasonality.

By integrating the relevant findings from Research Thematic I and systematically gaining the named information, a model is developed for all substations with a data collection rate of at least 90%. Using the findings established in Research Thematic II, fault-induced patterns in the resulting subsets of substation data are eliminated. The resulting model is evaluated to assess the hypothesis of the research presented in this documentation, highlighting the limitations and opportunities for application. For comparison, the heat demand calculated from the individual modeling targets is compared to the test data as well. A model with a 24-hour prediction horizon serves as a reference for comparison to existing literature and to assess the impact of this setting in contrast to nowcasts of the heat demand.

While substations with reliable, repetitive patterns can be modeled with comparably high success, fault-induced patterns and highly randomly distributed substation data require further investigation, particularly in the context of enhanced heat network control for efficiency and flexibility gains in highly dynamic heat network user data.

2.3.2 Publications related to Research Thematic III

- Submission in planning: WEBER, S./FISCHLSCHWEIGER, M./VOLTA, D.: Data pre-preparation, hyperparameter tuning, and domain-specific settings for dynamic user-level parameter modelling in heat network with deep learning LSTM.

3 Summary and Outlook

In recent years, there has been a growing interest among practitioners and researchers in increasing the efficiency of heat networks. The user as part of the heating system has increasingly become a subject of investigation for reasons of a continuous gain in data availability. The data open the possibility of applying data-driven approaches, such as ML. The thesis related to this documentation hypothesizes that user data are one critical resource for enhancing the efficiency and flexibility of heat networks, enabling the exploration of existing potentials within thermal energy systems. Therefore, the thesis associated with this documentation aims to identify effective combinations of ML architectures and available heat network data, achieving a high level of automation to generate insights based solely on accessible experimental data. This knowledge is intended to enhance the efficiency and flexibility of heat systems.

The thesis associated with this documentation contributes to the existing literature by providing systematic and methodological workflows for acquiring the knowledge necessary to develop and implement ML approaches in the context of thermal energy systems for a) modeling user-level parameters and b) effective fault management at the substation level within heat networks. The above-explained Research Thematics I to III present comprehensive approaches and findings, which are summarized as follows:

Research Thematic I: To implement the feature engineering and selection process for the newly suggested prediction targets, Research Thematic I establishes a systematic chain of methods. It also outlines the information on appropriate features for Research Thematic III. Generally, Research Thematic I highlights how the most sophisticated choice of features is individual to the selected modeling target. It presents knowledge of the user-level parameter modeling of heat networks, enabling the acceleration of the application of ML in heating systems.

Research Thematic II: Research Thematic II outlines that the primary side temperature data of substations enables fault detection with a high degree of automation. It showcases the application of k-means clustering to obtain suitable indicators for this purpose. The clustering method's objectivity enabled the identification of an indicator not yet widely applied, namely inverted temperature readings (supply temperature lower than return temperature). Research Thematic II showcases the successful encoding of the detected indicators for automated fault identification and impact quantification.

Research Thematic III: Using the features selected in Research Thematic I, Research Thematic III systematically tests settings for user-parameter modeling based on these. The results indicate that substations with repetitive, reliable behavioral patterns can be modeled successfully with the suggested setup. At the same time, those with high randomness in the collected readings or a strong change in the behavior between the training and test datasets could not be captured with comparable modeling performance. They were excluded using findings of Research Thematic II.

Optimizing operational parameters in heat networks necessitates predicting not only the optimal load to satisfy demand but also the minimum required temperature of the heat flow. Lowering network temperatures is crucial for achieving energy transformation within these systems. The thesis related to this documentation illustrates how user-level data can play a significant role in achieving these objectives. The findings presented offer valuable scientific and practical insights for advancing the use of ML in user-level data within heat networks, with the overarching aim of enhancing their efficiency and flexibility. The gained information can provide new knowledge for practical implementation and additional future research questions.

Three scientific publications have been prepared based on the research thematics: one published, one submitted for publication, and one prepared for submission. The author plans to finish the doctoral degree at the Clausthal University of Technology under the supervision of Prof. Fischlschweiger and Prof. Volta based on these research thematics. The author kindly thanks all supporters for their time, involvement, and interest, as well as any other means of assistance.

4 References

[1] 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, doi: 10.1109/IECON.2013.6700075, 2013.

[2] BAHLAWAN, 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.

[3] 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.

[4] 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.

[5] BUJALSKI, M./MADEJSKI, P./FUZOWSKI, K.: Day-ahead heat load forecasting during the off-season in the district heating system using Generalized Additive model. Energy and Buildings, 278, 112630, ISSN 0378-7788, https://doi.org/10.1016/j.enbuild.2022.112630, 2023.

[6] 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, 27th October 2022, Bundesamt für Justiz, 2021.

[7] 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.

[8] 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, Vol I, 2013.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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 2020, 3, Suppl. 1, 1-22, https://doi.org/10.1186/s42162-020-00125-5, 2020.

[15] 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, doi: 10.1109/TNNLS.2016.2582924, 2017.

[16] 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.

[17] JOHANSSON, C./BERGKVIST, M./GEYSEN, D./DE SOMER, O./LAVESSON, N./VANHOUDT, D.: 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, 2017.

[18] 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.

[19] KUDELA, L./CHÝLEK, R./POSPÍŠIL, J.: Efficient Integration of ML into District Heating Predictive Models. Energies, 13, 6381, 2020.

[20] 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.

[21] 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, 2020.

[22] 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, doi: 10.1109/ACCESS.2020.3017516, 2020.

[23] 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.

[24] 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.

[25] 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.jobe.2022.105316, 2022.

[26] 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.

[27] 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.

[28] MÅNSSON, 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.

[29] 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.

[30] NEUMAYER, M./STECHER, 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.

[31] 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.

[32] POTOČNIK, P./ŠKERL, P./GOVEKAR, E.: Machine-learning-based multi-step heat demand forecasting in a district heating system. Energy and Buildings, 233, 110673 https://doi.org/10.1016/j.enbuild.2020.110673, 2021.

[33] 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.

[34] 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, https://urn.kb.se/resolve?urn=urn:nbn:se:ltu:diva-40519, 2012.

[35] 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.

[36] SONG, J./ZHANG, L./XUE, G./MA, Y.P./GAO, S./JIANG, Q.L.: Predicting hourly heating load in a district heating system 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.

[37] 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.

[38] VESTERLUND, M./TOFFOLO, A./DAHL, J.: Optimization of multi-source complex district heating network, a case study. Energy, 126, 53-63, https://doi.org/10.1016/j.energy.2017.03.018, 2017.

[39] WEBER, S./FISCHLSCHWEIGER, M./VOLTA, D./RIECK-BLANKENBURG, U.: Identifikation und Einflussbewertung von Fehlern in Übergabestationen auf Endnutzerebene in Wärmenetzen. 2. Norddeutsche Wärmefachkonferenz, Hamburg, 19th – 20th September 2024.

[40] 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.

[41] 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.

[42] 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.

[43] 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.

[44] 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.

[45] 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.