Digital Twin-Based Fault Detection and Prioritisation in District Heating Systems

A Case Study in Denmark

Madsen, Frederik Wagner; Bank, Theis; Søndergaard, Henrik Alexander Nissen; Mortensen, Lasse Kappel; Shaker, Hamid Reza

Published in:
Energy Informatics

DOI:
10.1007/978-3-031-48652-4_18

Publication date:
2024

Document version:
Accepted manuscript

Citation for pulished version (APA):
Madsen, F. W., Bank, T., Søndergaard, H. A. N., Mortensen, L. K., & Shaker, H. R. (2024). Digital Twin-Based Fault Detection and Prioritisation in District Heating Systems: A Case Study in Denmark. In B. Nørregaard, L. C. Pereira da Silva, & Z. Ma (Eds.), *Energy Informatics : Third Energy Informatics Academy Conference, EI.A 2023, Campinas, Brazil, December 6–8, 2023, Proceedings, Part II* (Vol. 2, pp. 277–291). Springer. https://doi.org/10.1007/978-3-031-48652-4_18

Digital Twin-Based Fault Detection and Prioritisation in District Heating Systems: A Case Study in Denmark*

Frederik Wagner Madsen, Theis Bank, Henrik Alexander Nissen Søndergaard[0000−0002−9901−6728], Lasse Kappel Mortensen[0000−0002−4638−819X], and Hamid Reza Shaker[0000−0003−2858−8400]

University of Southern Denmark, Odense, Denmark

Abstract. Faults in district heating systems (DHS) cause sub-optimal operating conditions, which increase energy losses. As DHSs are critical infrastructure for many households in Denmark, these faults should be detected and corrected quickly. A novel model-based fault detection and diagnosis framework has been applied to detect and prioritise faults. The framework uses a bound for normal operation based on the residuals between historical sensor data and simulated properties in a digital twin of the DHS. The faults detected are prioritised based on the fault probability calculated using the Chernoff bound method. A case study on a Danish DHS has proven that the framework can produce a prioritised list of faults that maintenance crews can use to target faults with the highest probability. Furthermore, the digital twin allowed for fault location investigation, which could correlate different faults in the DHS. The framework has the potential for real-time fault detection and diagnosis. However, more precise digital twins need to be developed.

Keywords: fault detection and diagnosis · district heating systems · digital twin · Chernoff bound.

1 Introduction

A district heating system (DHS) aims to distribute and provide affordable heat to connected consumers efficiently [4]. In Denmark, the majority of households' heat is supplied by district heating (DH), and a large proportion of total energy use is for hot water use and heating in the EU [6], which makes the DHSs critical infrastructure, and an important domain to improve the energy efficiency of, to lower CO_2 emissions. The complexity of operating a DHS with all the different functions, from heat production and transmission to consumption, leads to faults. These faults result in the DHS operating sub-optimally [4, 11, 12]. To ensure the DHSs work efficiently, it is important to detect these faults and remedy

* This work is supported by the "Proactive and Predictive Maintenance of District Heating Systems" and "IEA DHC TS4", funded by the Danish Energy Agency under the Energy Technology Development and Demonstration Program, ID number 64020-2102 and 134-22011, respectively.

them as quickly as possible. However, DH companies mostly perform reactive- and preventive maintenance instead of proactive maintenance. Reactive maintenance strategies often result in faults not being detected quickly enough or are never detected if the system is able to compensate or does not affect the delivery of heat to the consumers (but still wastes energy). Faults leading to the discomfort of consumers are often detected far more quickly. The use of planned maintenance results in inspections being done too often, resulting in a waste of resources. To unleash the full potential of the digitalisation of the DH sector, tools for enabling proactive maintenance via implementing fault detection and diagnosis (FDD) methods are needed. These methods must be able to detect faults in a DHS in a timely and economical manner to reduce waste of energy and time and lowering CO_2 emissions.

Methods for FDD are categorised in [8] into three sub-types: quantitative model-based, qualitative model-based, and process history-based, which all come with advantages and disadvantages. Quantitative model-based, which uses a model based on thorough physical or engineering principles, enhances the precision of the model. However, comprehensive modelling is also a weakness due to the level of complexity and amount of input required to model, which can reduce the scalability of the approach. The method is used in [19], where a DHS is modelled in OpenModelica. The residuals between model output and pressure sensors are compared in a Bayesian Network to determine the possibility of faults and evaluate the system. Another example is [2], in which a simulation of a DHS grid is also created and is coupled with an optimisation problem to detect and identify both thermal and hydraulic faults.

Both quantitative and qualitative model-based FDD have some overlapping elements, and the differences can sometimes be vague. On the other hand, process history-based FDD differentiates itself by only utilising data. This makes the modelling less complex because black-box models require less knowledge of the physics of the system and make it easier to scale to other systems. However, the method requires a lot of good-quality data and, to some extent, computational power. An example of that is the method seen in [15], where three methods of FDD, Hotelling's T^2 and Q statistics, contextual Shewhart chart, and linear regression, are presented. Furthermore, [15] uses an approximation of the Chernoff bound method proposed in [5] to investigate if all three FDD methods agree on the same fault thereby filtering out insignificant alarms for faults occurring in the DHS. Based on the identified literature on FDD in the DHS domain, it is apparent that more research employs data-driven methods compared to model-based methods. Some examples of data-driven research are Sun et al. [14], where they use three clustering methods to identify operation patterns that lead to faulty behaviour in consumers. The use of a gradient boost regressor has been utilised by Månsson et al. [10] to detect faults in DH substations by predicting regular operation. Lastly, the work in [17] utilises cluster analysis and association analysis- to decide rule patterns for the operation of the DH substation. According to a review paper on FDD in DHS [4], many of the data-driven models created will not perform well and are thereby not useful in the real world, due to them being created and trained using laboratory or simulation data. This problem can be alleviated either by using real-world data or by utilising model-based methods. In general, the review paper [4] indicates an overall research gap in the field of FDD in DHSs. This can be in part due to the preconceived idea that DHS work well, even though that is not the case [7], and the use of DH is not as widespread.

A challenge emphasised by the literature is the number of faults different FDD methods will produce [9, 13]. A methodology to differentiate between different faults more accurately is important in the decision-making process for making corrective maintenance in DHSs for liability and economic reasons. Paper [19] compares the model's output with sensor data from the system and detects a fault when the sensor data deviates $\pm 1\%$ from the model simulation in at least ten consecutive time steps. The fault is then run through a Bayesian Network to diagnose the fault. The Bayesian Network is built upon expert knowledge, strengthening the capability of diagnosing faults. Still, on the other hand, expert knowledge may not be sufficient to identify a correlation between residuals to a fault diagnosis, thereby disregarding a possible fault. The current approach to fault prioritisation is mainly based on expert knowledge to classify the severity of faults and this is extremely time-consuming [16]. Furthermore, according to [16], data-driven methods are not widely utilised for fault prioritisation or even seen in literature at all. Some developments have been made, however. [5] proposed a statistical method of prioritising faults in a telecommunication network using the Chernoff bound method. Using the Chernoff bound method, this statistical approach can identify the probability of a fault occurring in a system without expert knowledge. This methodology has been brought into the energy domain in papers [1,3] using it for FDD in a heat, ventilation, and air conditioning (HVAC) system of a building. Using the Chernoff bound method, the papers were able to categorise the state of urgency into three subsets: low($P < 60\%$), medium($60\% \leq P < 95\%$), and high($P \geq 95\%$). The model-based FDD using the Chernoff bound method proposed in the papers [1, 3] showed great potential for FDD in HVAC systems but also the potential for implementation in other energy-system domains such as DHSs.

This paper proposes an FDD framework for DHSs. In summary, it applies Chernoff bound methodology to the residuals between the outputs of a quantitative model of a DH network (digital twin) and real-world sensor data, when the residual is above a predetermined threshold. Chernoff bound provides a fault probability value that can be used by maintenance for better prioritisation of which faults to further investigate. This paper aims to contribute with a framework to the apparent research gap in the literature for FDD in DHSs with a two-fold contribution.

1. A framework is developed by implementing a quantitative model-based FDD approach together with the Chernoff bound method to prioritise faults by

their probability. This makes it possible for maintenance crews to only focus on the faults with the highest fault probability, not wasting their time. To the best of our understanding, this has not been done before.
2. The framework's methodology is tested using a case study with real-world data and a digital twin of the DHS.

In section 2, the overall framework for model-based FDD in DHSs is introduced and explained in detail. The next step is in section 3, in which the framework is applied to a case study, and results are presented.

2 Methodology

A novel model-based FDD framework will be presented in this section. The section will describe how sensor data measured in a DHS is processed and simulated in a digital twin for FDD and how the faults are indexed according to their probability. A flow diagram of the framework can be seen in Fig. 1. For the purpose of this work, we adopt the general definition of a digital twin stated by Yu et al. [18] that "a digital twin is a digital (or virtual) representation that looks-like, behaves-like, and connects-to a physical part or system with the goal of improving or optimising decision making for any time horizon".

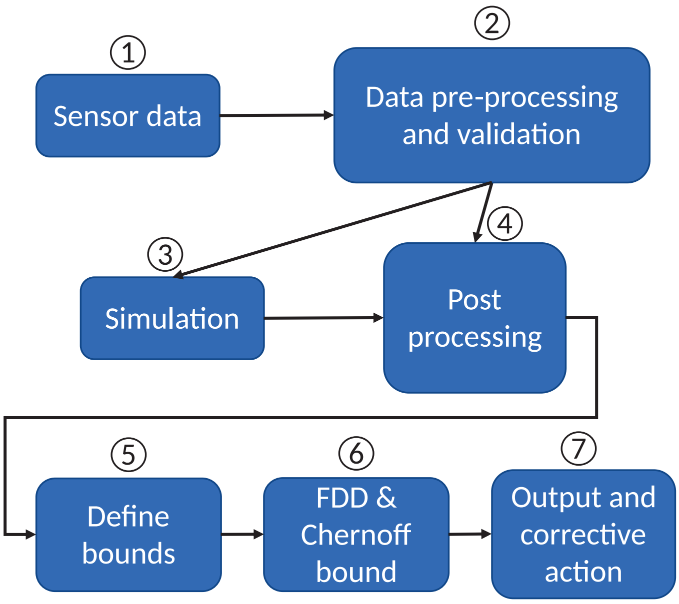

Fig. 1. Flow diagram of model-based FDD framework.

In step 1, the data is collected from the sensors in the DHS. The sensor data is pre-processed to have the right format and units to match the requirements for the data input to the digital twin in step 2. The data validation in step 2 investigates if negative- or NAN-values are present in the data. The data validation also investigates if some measurements violate the physical capabilities of the system or if unexpected repeating patterns can be found in the data, which could indicate sensor faults. Invalidated data must be corrected before further use of the data. In step 3, a digital twin uses some of the data collected by the sensors as boundary conditions to perform quasi-dynamic simulations of the state of the DHS. The digital twin should be a very accurate digital model of the DHS where the pipe configurations, heat loss coefficients etc., are defined. The results from the digital twin are then post-processed together with the unused sensor data for FDD. This model-based FDD framework proposes a univariate statistical approach using a bound of normal operation to detect faults. The bound of normal operation is defined as the root-mean-square error (RMSE) of the residual of one property, e.g., mass flow rate, between the digital twin and the sensor data. With this approach, it is assumed that the majority of the sensors in the DHS are correct and if one sensor deviates more than the bound of normal operation, it is detected as a fault. Furthermore, the bound also represents the modelling error of the digital twin, which allows some deviations from the norm. The RMSE value is calculated using Equation 1, where N and M are the time steps and sensors, respectively.

$$RMSE = \sqrt{\frac{\sum_{k=1}^{M}\sum_{i=1}^{N}(Y_{ik}^{sensor} - Y_{ik}^{digital\ twin})^2}{N \cdot M}} \tag{1}$$

The RMSE can be interpreted as the standard deviation of the residuals. The r-value, which is the number of standard deviations included, can adjust the size of the bound. The FDD should not raise an alarm if the sensor measurement is within the bound, as seen in Equation 2.

$$\begin{aligned} &Y_i^{sensor} > Y_i^{digital\ twin} - r \cdot RMSE \\ \text{and}\quad &Y_i^{sensor} < Y_i^{digital\ twin} + r \cdot RMSE \end{aligned} \tag{2}$$

Due to some data having higher numerical values, the data is normalised with maximum scaling (see Equation 3) before the RMSE value is calculated. The maximum scaling is done by finding the maximum value in the data for one consumer k, which has one column with sensor data and one with simulated data. For each time step i, the data point $r_{i,k}$ is divided by the maximum value $max(r_k)$ giving the normalised data point $n_{i,k}$. Maximum scaling ensures that all data is less than or equal to 1.

$$n_{i,k} = \frac{r_{i,k}}{max(r_k)} \tag{3}$$

In step 6, FDD is carried out using the bound of normal operation. To prioritise which faults have the highest probability and should be investigated by a maintenance crew, the Chernoff bound method from [5] is used. A schematic of the Chernoff bound method can be seen in Fig. 2.

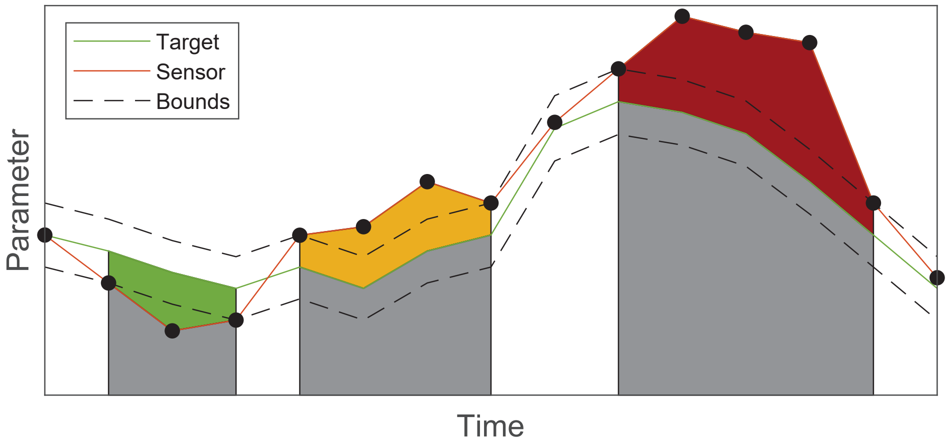

Fig. 2. Chernoff bound schematic. Showcasing the gray and coloured areas used for calculating the probability. The figure does not depict real data.

A suspicion is started when the sensor measurements leave the bound of normal operation and end when it enters again. The areas under the curves are used in Equations 4 and 5 to calculate the probability of normal- and faulty operation, respectively. As the areas A_O and A_B are used to calculate the probability, the fault probability for a period is correlated to the amplitude and the period.

$$P_{normal\ operation} = P(A_O|A_B) = e^{-\left(\frac{\left(A_B\left(1-\frac{A_O}{A_B}\right)^2\right)}{2}\right)} \tag{4}$$

$$\begin{aligned} P_{faulty\ operation} &= 1 - P_{normal\ operation} \\ A_O &= Grey\ area \\ A_B &= Grey\ area + Colored\ area \end{aligned} \tag{5}$$

Lastly, in step 7, the output of the Chernoff bound method is prioritised from highest fault probability to lowest. As proposed in papers [1, 3], the faults are categorised into three fault probability indices; high, medium, and low, as seen in table 1.

Table 1. Fault probability indexing levels classifications.

Probability level index	Probability of fault	Colour indication
High	$P \geq 95\%$	Red
Medium	$60\% \leq P < 95\%$	Yellow
Low	$P < 60\%$	Green

3 Case Study and Results

3.1 Description of Case Study

The model-based FDD framework presented will be implemented in a case study with historical data from a suburb of the DHS in Odense, Denmark. The investigation period is from December 2022 to January 2023. A sensor installed at the substation measures supplied and returned energy, mass flow rate, pressure, and temperature at an hourly resolution. At every 648 consumers, a sensor measures the volumetric flow rate, energy consumption, and supplied and returned energy at a daily resolution. Supply and return energy is the supply and return temperature multiplied by the volumetric flow rate. The digital twin of the DHS, developed by a collaboration between Fjernvarme Fyn and Danfoss, was built in the software Leanheat Network (LHN) and will be used to simulate the DHS. LHN simulates hydraulic and thermal conditions based on boundary conditions and optimises pressure-, mass flow rate-, and temperature conditions to minimize pump- and heat production costs. The digital twin will simulate hourly quasi-dynamic simulations, where each hourly simulation will represent the whole day. The boundary conditions imported to the digital twin are the power consumption and return temperature at the consumers, supply temperature and return pressure at the substation, and the pressure change at a critical node defined as 1.33 bar, based on Fjernvarme Fyn's expert knowledge. The sensor data is reformatted to meet the data import requirements for the software. The output from the digital twin, which will be used for FDD, is the supply temperature and mass flow rate at the consumers. Implementing the Chernoff bound method will produce a list of fault probabilities for these two properties with the format as seen in table 2.

Table 2. Fault probability list of mass flow rate FDD.

Consumer ID	Start	End	Fault probability	Fault probability index
Consumer_345	2023-01-04	2023-01-16	98%	High
Consumer_322	2022-12-07	2022-12-14	84%	Medium
⋮	⋮	⋮	⋮	⋮
Consumer_103	2023-01-05	2023-01-06	16%	Low
Consumer_382	2022-12-15	2022-12-16	16%	Low

3.2 Calibration of Bound Size

As the bound size is defined by $Y_i^{digital\ twin} \pm r \cdot RMSE$, the number of faults found and their probability are directly correlated with the chosen r-value. The investigation of every single fault can be costly and time-consuming. Therefore, tuning the r-value to ensure that a reasonable number of faults with high probability are found is important, disregarding the faults with low probability caused by, e.g., the modelling error. This paper suggests tuning the r-value manually by investigating the framework's output when the r-value is changed. Another approach for automatically tuning the r-value could be the elbow method which is a heuristic method for finding the incremental increase in the r-value with the largest marginal decrease in the number of faults detected. However, the elbow method was not investigated as it is seen as more beneficial for the operator to manually tune the number of faults detected to match the resources available to investigate them. Automatically tuning the r-value could produce an unnecessary amount of faults detected which can be unmanageable to investigate. The proposed manual tuning of the r-value is visualised in Fig. 3 for the mass flow rate and Fig. 4 for the supply temperature, where the number of faults and mean probabilities are shown for each r-value.

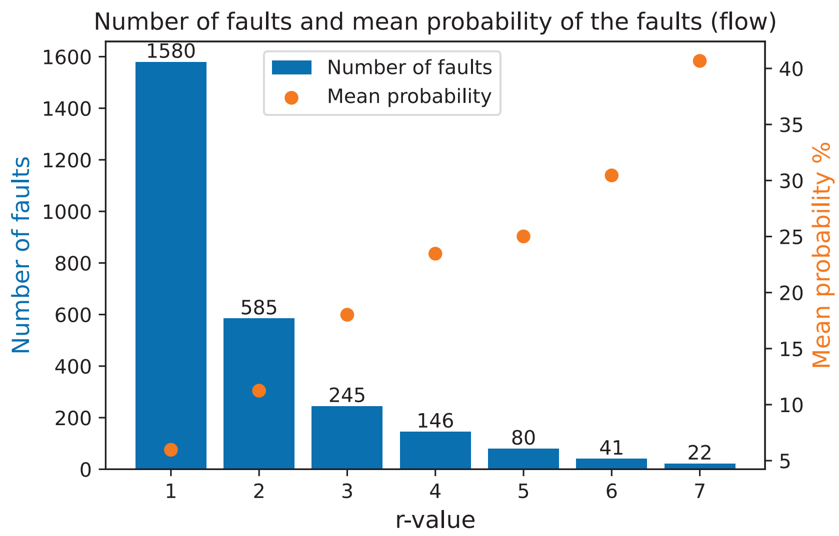

Fig. 3. Incrementally increasing the r-value from one to seven for the mass flow rate FDD.

Fig. 3 shows that the number of detected faults decreases with increasing r-value, contrary to the mean probability of the detected faults, which increases with increasing r-value. This illustrates how the size of the bound and the Chernoff bound method are combined only to detect and prioritise the most important faults. For further fault investigation, an r-value of 7 is chosen, which gives 22 faults (13 unique consumers) with a mean probability of 40%.

Looking at supply temperature, the investigation of the r-value shows similar results, seen in Fig. 4. For the further investigation of supply temperature, an r-value of 5 has been chosen, which resulted in 12 faults with five unique consumers and a mean probability of around 50%. R-values larger than 5 seem to decrease the number of faults at a low rate, and it is therefore not seen as necessary to increase the r-value above 5. Another approach is to choose a low r-value and then sort the list of faults, only looking at the high-probability faults.

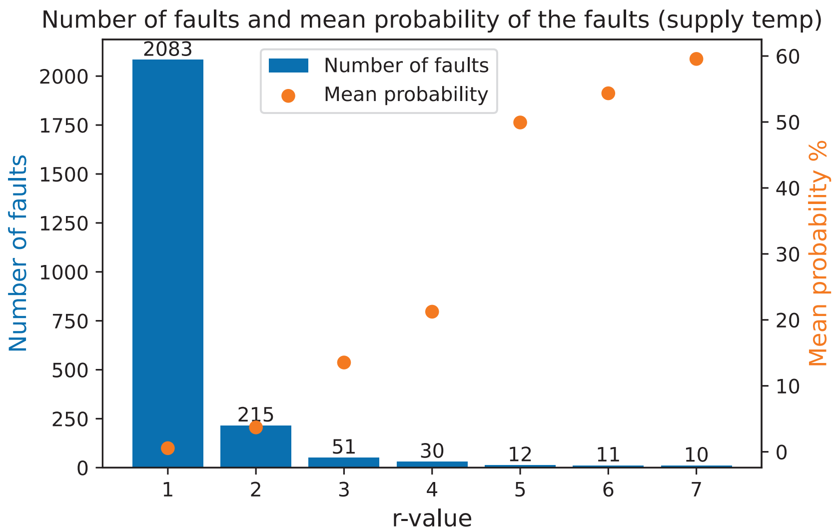

Fig. 4. Incrementally increasing the r-value from one to seven for the mass flow rate FDD.

3.3 Fault Investigation

In Fig. 5, Consumer_239 is showcased, which operates under normal conditions. It can be seen that the sensor- and simulated boundary conditions are equal. Some deviations exist between the measured- and simulated mass flow rate and supply temperature. Still, these small deviations were inside the bound of normal operation and were therefore not detected as faults. The bound of normal operation is not shown in Fig. 5, as the bound was calculated based on normalised data.

From the prioritised list of fault probabilities, regarding the supply temperature FDD, consumer_471 experienced three faults. The operation of consumer_471 can be seen in Fig. 6, and the three faults (grey areas) had a fault probability of 35%, 95%, and 58%, respectively. Correlating the four properties in Fig. 6, the largest deviation in supply and return temperature happens in periods when consumer_471 has no consumption. When there is no consumption, the mass flow rate is also 0 kg/s, which may cause the sensor not to measure

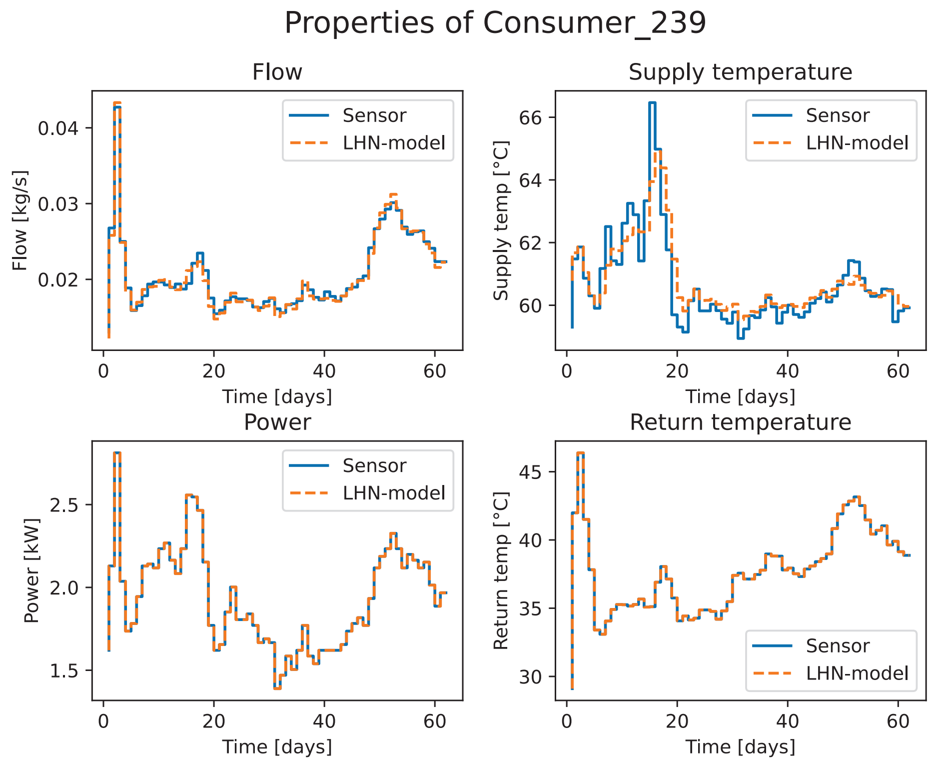

Fig. 5. Consumer_239 shows normal operation conditions, as the residuals between the sensor- and model curves for the four properties are within the bound of normal operation.

the supply and return temperature correctly and set it to 0 $°C$. These boundary conditions can not be simulated accurately by the digital twin, which forces it to change the boundary condition causing this large deviation. This fault may be labelled as a sensor fault and will probably not have damaging effects on the system. However, it can be noticed that the digital twin also changes the return temperature, which is a boundary condition, after the period of no consumption ends by December 2022. In the global property settings in the LHN software, it was defined that the simulation will not allow a $\Delta T \leq 0.5$ $°C$, i.e., that the consumers cool the DH water less than 0.5 $°C$. For consumer_471 $\Delta T \geq 0.5$ $°C$, however, the digital twin still changes the defined return temperature. This action by the digital twin, where boundary conditions were changed, is due to numerical stability in the optimisation problem being solved (confirmed by Danfoss).

Consumer_365, presented in Fig. 7, had two faults in the investigation period with 26% and 48% fault probability, respectively, regarding the FDD on

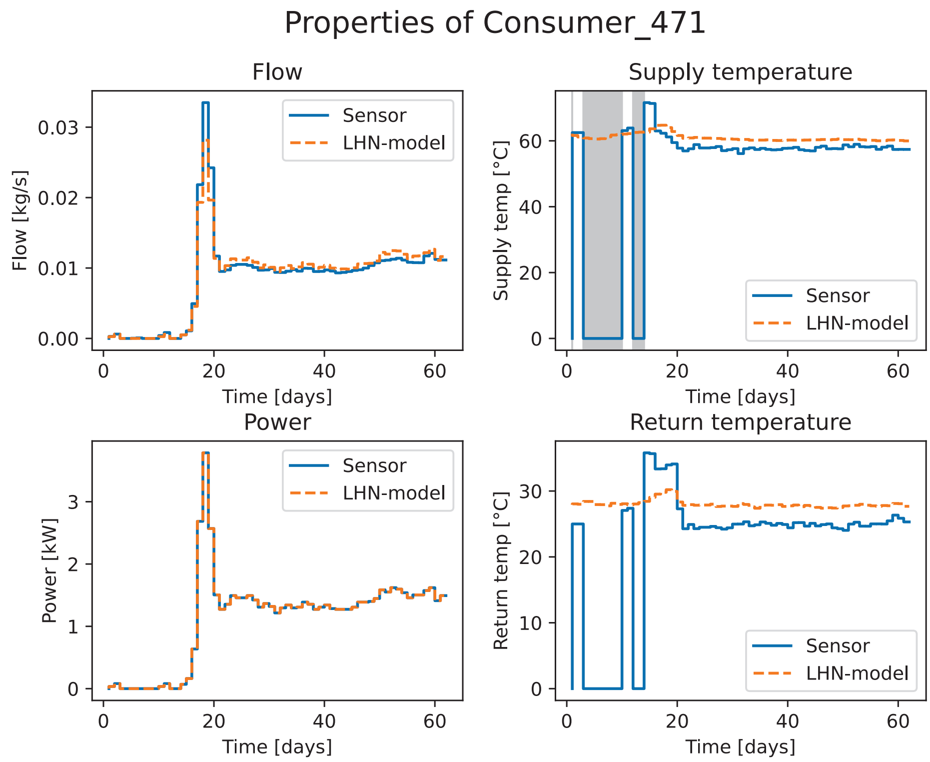

Fig. 6. Consumer_471 shows faulty operation, where the marked grey areas for the supply temperature chart, within the first 15 days, are the three faults.

mass flow rate. These faults did not rank high on the fault probability indexing. Nevertheless, Consumer_365 is showcased as Fjernvarme Fyn validated that consumer_365 was operating in a faulty condition due to a low cooling efficiency of the consumer installation and that these faults showed particular interest compared to higher prioritised faults. The faults can be seen in Fig. 7, where the simulated mass flow rate spikes twice in the investigation period (grey areas). These spikes are infeasible to occur in the real DHS. The digital twin calculates results by running an optimisation problem, where it tries to match all properties in the pipes and nodes in the system given the set of boundary conditions. These spikes indicate that the boundary conditions for consumer_365 could not be simulated accurately by the digital twin and therefore have a probability of being faulty.

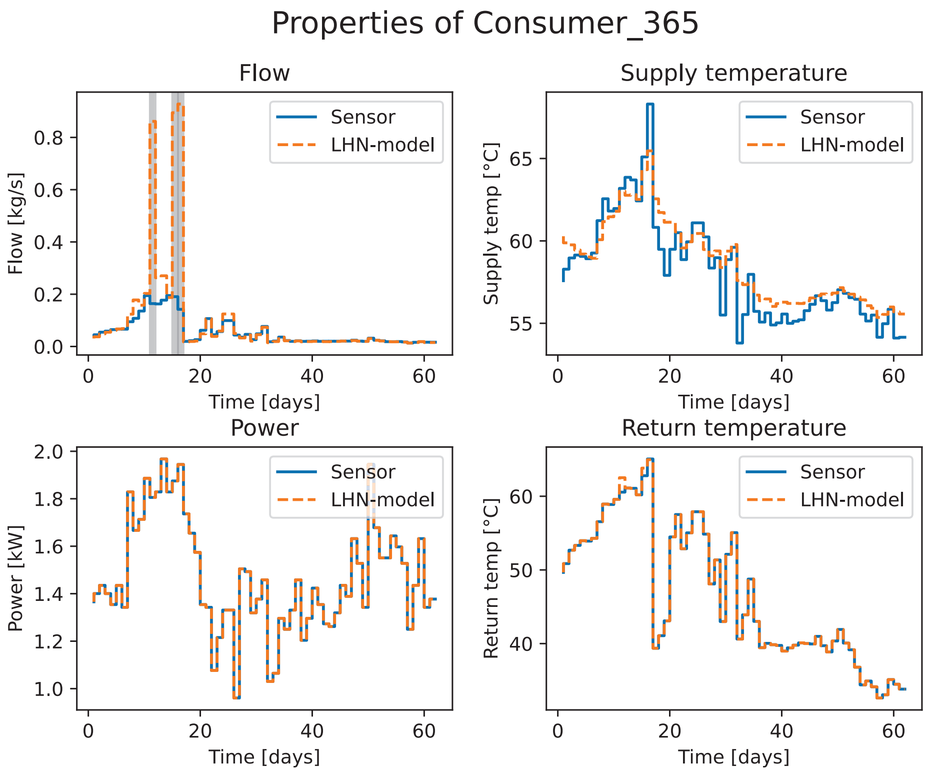

Fig. 7. Consumer_365 shows a faulty operation, where the marked grey areas for the mass flow chart are the two faults.

3.4 Fault Location Investigation

Instead of investigating each fault individually on the list of faults, another approach is to identify the consumers' location and the types of houses on the list. For this investigation, having a digital twin with a geographical user interface is advantageous. Using the two lists of faults obtained by an r-value of 7 for mass flow rate and 5 for supply temperature, the fault location investigation shows that all the detected faults happened in terrace houses. The DHS in the suburb has more terrace houses than single houses, but this indicates that the terrace houses are more prone to operating in faulty conditions. Another interesting result is the location of three out of four consumers in the same terrace house had a fault detected, illustrated in Fig. 8.

The terrace house is located far from the substation and is also the last connection point of that pipeline branch, where the non-faulty consumer is the first one connected. Using this knowledge of the faults' locations can indicate the faults' cause, as similar faults may unlikely occur at three of the four consumers.

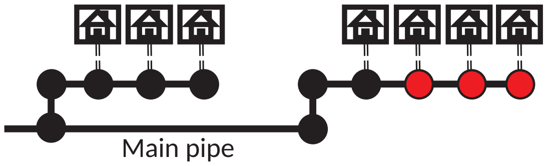

Fig. 8. Three out of four consumers in same terrace house operating faulty.

This could indicate a fault upstream in the pipe network, possibly the pipe going from the main pipe into the terrace house. However, a thorough investigation of the pipes and consumers must be done to find the root of the three faults occurring in the terrace house.

4 Conclusion and Future Work

Faults in district heating systems (DHS) that are not detected quickly enough or not at all, can waste a lot of energy. As DHSs are critical infrastructure for the many Danish households connected, fault detection and diagnosis (FDD) frameworks are of great importance in detecting and correcting these faults. A model-based FDD framework detecting and prioritising faults in a DHS has therefore been developed in this paper. The model-based FDD framework detects faults using a bound of normal operation based on the root-mean-square error (RMSE) of the deviation between a digital twin and the sensors in a DHS, where the r-value could calibrate the number of faults detected, disregarding faults caused by the modelling error. The faults detected were prioritised based on their fault probability using the Chernoff bound method. The prioritised list of faults can be useful for maintenance crews as they can save time and resources by investigating and targeting faults with a high probability. The framework is scalable and can easily be implemented by DH companies with existing digital twins. For future work, the model-based FDD framework should be tested on live sensors, where the benefits of enhanced time resolution of the sensor measurements also could be investigated.

Acknowledgements We thank Peer Andersen, Lasse Elmelund Pedersen, and the rest of their team at Fjernvarme Fyn A/S for their assistance with the data and the model. Also, thanks to Johan Peter Alsing from Danfoss A/S for assisting us with Leanheat Network.

References

1. Alexandersen, E.K., Skydt, M.R., Engelsgaard, S.S., Bang, M., Jradi, M., Shaker, H.R.: A stair-step probabilistic approach for automatic anomaly detection in building ventilation system operation. Building and Environment (2019). https://doi.org/10.1016/j.buildenv.2019.04.036

2. Bahlawan, H., Ferraro, N., Gambarotta, A., Losi, E., Manservigi, L., Morini, M., Saletti, C., Ruggero Spina, P., Venturini, M.: Detection and identification of faults in a district heating network. Energy Conversion and Management **266**, 115837 (2022). https://doi.org/10.1016/j.enconman.2022.115837
3. Bang, M., Engelsgaard, S.S., Alexandersen, E.K., Skydt, M.R., Shaker, H.R., Jradi, M.: Novel real-time model-based fault detection method for automatic identification of abnormal energy performance in building ventilation units. Energy and Buildings (2019). https://doi.org/10.1016/j.enbuild.2018.11.006
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. Applied Sciences **11**(1) (2021). https://doi.org/10.3390/app11010455
5. Cheung, B., Kumar, G., Rao, S.A.: Statistical algorithms in fault detection and prediction: Toward a healthier network. Bell Labs Technical Journal **9**(4), 171–185 (2005). https://doi.org/10.1002/bltj.20070
6. EU: Commission recommendation (eu) 2019/786 of 8 may 2019 on building renovation (2019), https://bit.ly/30nxBs5
7. Gadd, H., Werner, S.: Fault detection in district heating substations. Applied Energy **157**, 51–59 (2015). https://doi.org/10.1016/j.apenergy.2015.07.061
8. Katipamula, S., Brambley, M.R.: Review article: Methods for fault detection, diagnostics, and prognostics for building systems—a review, part ii. HVAC&R Research **11**(2), 169–187 (2005). https://doi.org/10.1080/10789669.2005.10391133
9. Kim, W., Katipamula, S.: A review of fault detection and diagnostics methods for building systems. Science and Technology for the Built Environment **24**(1), 3–21 (2018). https://doi.org/10.1080/23744731.2017.1318008
10. Månsson, S., Kallioniemi, P.O.J., 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, 16th International Symposium on District Heating and Cooling, DHC2018, 9–12 September 2018, Hamburg, Germany
11. Pakanen, J., Hyvärinen, J., Kuismin, J., Ahonen, M.: Fault diagnosis methods for district heating substations. VTT Tiedotteita - Valtion Teknillinen Tutkimuskeskus (01 1996)
12. Sandin, F., Gustafsson, J., Delsing, J.: Fault detection with hourly district energy data: Probabilistic methods and heuristics for automated detection and ranking of anomalies. Tech. rep., Svensk Fjärrvärme AB (2013)
13. Shaker, H.R., Santos, A., Jørgensen, B.: A practical data-driven condition indicator for room-level building diagnostics. Energy Informatics (09 2021)
14. Sun, W., Cheng, D., Peng, W.: Anomaly detection analysis for district heating apartments. Journal of Applied Science and Engineering **21**, 33–44 (3 2018). https://doi.org/10.6180/jase.201803_21(1).0005
15. Søndergaard, H.A.N., Shaker, H.R., Jørgensen, B.N.: Automated and real-time anomaly indexing for district heating maintenance decision support system (preprint). SSRN Electronic Journal (2023). https://doi.org/10.2139/ssrn.4344182
16. Webert, H., Döß, T., Kaupp, L., Simons, S.: Fault handling in industry 4.0: Definition, process and applications. Sensors **22**(6) (2022). https://doi.org/10.3390/s22062205, https://www.mdpi.com/1424-8220/22/6/2205
17. Xue, P., Zhou, Z., Fang, X., Chen, X., Liu, L., Liu, Y., Liu, J.: Fault detection and operation optimization in district heating substations

based on data mining techniques. Applied Energy **205**, 926–940 (2017). https://doi.org/10.1016/j.apenergy.2017.08.035
18. Yu, W., Patros, P., Young, B., Klinac, E., Walmsley, T.G.: Energy digital twin technology for industrial energy management: Classification, challenges and future. Renewable and Sustainable Energy Reviews **161**, 112407 (2022). https://doi.org/https://doi.org/10.1016/j.rser.2022.112407, https://www.sciencedirect.com/science/article/pii/S136403212200315X
19. Zimmerman, N., Dahlquist, E., Kyprianidis, K.: Towards on-line fault detection and diagnostics in district heating systems. Energy Procedia **105**, 1960–1966 (2017). https://doi.org/10.1016/j.egypro.2017.03.567, 8th International Conference on Applied Energy, ICAE2016, 8-11 October 2016, Beijing, China