*Manuscript

Large-scale monitoring of operationally diverse district heating substations: A reference-group based approach

Shiraz Farouq[a,*], Stefan Byttner[a], Mohamed-Rafik Bouguelia[a], Natasa Nord[b] and Henrik Gadd[c]

[a]Department of Intelligent Systems and Digital Design, Halmstad University, Sweden
[b]Department of Energy and Process Engineering, Norwegian University of Science and Technology, Trondheim, Norway
[c]Öresundskraft, Helsingborg, Sweden

ARTICLE INFO

Keywords:
district heating substations
return temperature
reference-group based operational monitoring
fault detection
outlier detection

ABSTRACT

A typical district heating (DH) network consists of hundreds, sometimes thousands, of substations. In the absence of a well-understood prior model or data labels about a substation, the overall monitoring of such large number of substations can be challenging. To overcome the challenge, an approach based on collective operational monitoring by a local group (i.e., the reference-group) of other similar substations in the network was formulated. Herein, if a substation of interest (i.e., the target) starts to behave differently in comparison to those in its reference-group, then it was designated as an outlier. The approach was demonstrated on the monitoring of the return temperature variable for atypical[1] and faulty operational behavior in 778 substations associated with multi-dwelling buildings. The choice of an appropriate similarity measure along with its size k were the two important factors that enables a reference-group to detect outliers in an effective manner. Thus, different similarity measures and size k for the construction of the reference-group were investigated. This led to the selection of Euclidean distance as a similarity measure with $k = 80$. This setup resulted in the detection of 44 target substations that were outliers, i.e., the behavior of their return temperature changed in comparison to the majority of those in their respective reference-groups. In addition, six frequent patterns of deviating behavior in the return temperature of substations were identified using the reference-group based approach, which were then further corroborated by the feedback from a DH domain expert.

1. Introduction

The survival of DH industry in the future will rely on its ability to use a wide range of sustainable energy sources like biomass, geothermal, industrial excess heat and domestic and industrial waste. This would require efficiency in both the supply and the demand side of energy use. The most potent way to move towards this goal is to reduce the distribution temperature of the DH networks. However, the current situation in Sweden is that DH networks have supply temperatures of about $75-90°C$ and return temperatures of about $40-50°C$ as annual averages (Frederiksen and Werner, 2013). One major reason behind a high supply temperature is that faults, both at the primary and secondary side of a district heating substation[1], are compensated through the supply temperature increase. However, this situation cannot be sustained if a transition is to be made to the 4^{th} generation district heating (4GDH) technology (Lund et al., 2014), where low distribution temperatures of $50/20°C$ are a key requirement. In a 4GDH regime, DH utilities will not be able to compensate for faults in their respective DH networks through a high supply temperature, which in turn will directly affect their customer's comfort. It has been shown in (Gummérus, 1989) that if all substations work as designed, the current distribution temperatures of DH networks can be decreased to approximately $70/35°C$. Moreover, according to (Sköldberg and Rydén, 2014), a decrease in the return temperature of the DH networks can result in a cost saving of up to 1 billion SEK[2] per year for the DH sector in Sweden. This clearly indicates that monitoring and fault detection capabilities must be enhanced at the substation level of the DH network.

An adequate way to deal with the monitoring of substations is to identify a physical model for each building's thermodynamics in conjunction with its heating system (Bacher and Madsen, 2011), while also taking into account social factors (Yao et al., 2009). These models can then be used for various purposes, including the control of indoor

[1]Here, "atypical" means that while it does not fit the definition of a normal operation, it is not faulty either and may also have some context.

*Corresponding author
shiraz.farouq@hh.se (S. Farouq)
ORCID(s):

[1]From now on, the word "substation" will be used to refer to a district heating substation.
[2]150 million USD (1 SEK≈0.15USD (annual average exchange rate) in 2014).

climate, forecasting of thermal energy consumption, description of building's energy performance and fault detection. However, such models can require a lot of details about each building and its heating system, which makes them costly and difficult to construct. Therefore, construction of such models, especially on large-scale, is mostly infeasible. A similar observation has been made in (Gustafsson and Sandin, 2016):

A substation and the associated building is a complex physical system that is difficult to model because of the high number of substations in a DH system.

Due to the aforementioned reasons, no prior benchmark reference models are available in practice for the monitoring of buildings and their associated heating systems.

The energy policy of Sweden shares a common basis with energy policy developed at the European Union (EU) level. In 2006, the EU adopted the energy end-use efficiency directive (European Parliament, Council of the European Union, 2006). It suggests improving billing by taking measures to charge the consumer in a timely manner based on their actual energy use. It also calls to enable final consumers to make better informed decisions as regards their individual energy consumption by providing them with further relevant information on energy use. As a result, the Swedish DH act (Riksdagen, 2008) was changed, requiring that from 1^{st} January 2015, all DH entities to charge their customers based on their actual monthly thermal heat use. Therefore, substations are now equipped with digital metering devices on the primary side, measuring hourly heating rate, flow rate, the supply temperature and the return temperature. These meters are, however, only data acquisition (DAQ) systems. The main use of this data acquisition for the DH utilities is to bill the customers for heat use.

One can expect that the data available through digital meters enables the use of data-driven approaches to monitor and study the operational behavior of substations in a DH network. However, the main assumption of a data-driven approach is that all relevant information about a system is sufficiently well-captured in the data. Most commonly, three possible schemes exist here: (a) a data-driven model for each substation, (b) a hybrid model for each substation, and (c) a global model for the entire DH network.

The first scheme has limitations due to the non availability of important information, such as, data related to the control parameter setting and secondary side operation of the substations. Moreover, the available data lacks labels on normal, atypical and faulty operational behavior of a substation. One reason why the aforementioned data and details related to substations are not available to a DH utility is due to the issue of ownership. In most cases in Sweden, it is the building administration or the house owner that owns a substation and not the DH utility. Right now, DH utilities in Sweden do not have a procedure in place to obtain details about a substation or its history of fault complaints and subsequent repairs. Furthermore, the operational behavior of substations in a network can vary depending on various factors. These include local weather conditions and different characteristics of the buildings, such as, geographical location, construction year and thermal transmittance value (U-value). Additional factors include the social purpose of the buildings and its associated control strategy, which then affects its heat load profile (Frederiksen and Werner, 2013; Gadd and Werner, 2013). Details related to these factors are also usually not available. Additionally, tuning for various hyper-parameters of a selected data-driven model for each substation is required. Hence, a data-driven model for each substation can be difficult to obtain and scale to all the substations in a large DH network.

The second scheme is based on a hybrid approach which combines physical models with data-driven models, while also incorporating expert opinion. For instance, in (Cai et al., 2019), a dynamic Bayesian network (DBN) model is proposed. Herein, when data on certain variables is not available for the construction of a DBN model of a system, physical models are used to estimate their distribution or the value of their parameters. In some cases, their values were obtained based on expert opinion. However, buildings and their associated heating systems in a city are quite diverse. Therefore, input parameter values for even simple physical models may not be readily available.

To mitigate for the non availability of data and various details related to substations to some extent, the third scheme provides for an alternative approach, which assumes that the operational behavior of substations in a DH network is homogeneous. This implicitly assumes that substations are affected by the same set of unobservable variables over time. This approach is referred to as group or fleet based monitoring, and requires a global model for the entire fleet or network[3] (Byttner et al., 2011; Oza and Das, 2012). Such a setup would require each substation to be described by the same set of representative features which are then used as input to an appropriate data-driven model for the DH network. Any substation that deviates significantly from the network according to this model is considered as atypical or faulty. Such global models can be inefficient in detecting atypical or faulty substations. This is because, in most cases, substations exhibit operationally diverse behaviors due to both technical and social factors.

[3] For referring to large number of substations, the word "network" is more appropriate than the word "fleet".

In summary, typical data-driven schemes for operational monitoring of substations are constrained by the following factors:

1. Lack of access to data and information related to the buildings and their heating systems.
2. Absence of labels about atypical and faulty behavior together with their associated contexts.
3. Operational diversity due to social and technical factors.

In general, the most common state of practice in the DH industry is to use thresholds (Månsson et al., 2019). The main limitation of using such a method is the choice of threshold values. In the presence of operational diversity, achieving a compromise between efficiency of detecting true outlier substations and false alarms on the basis of a threshold is mostly unattainable.

The objective of this study is to address the aforementioned constraints on the use of typical data-driven schemes and the limitations of the state of practice for the operational monitoring of a large network of substations. To achieve this, a reference-group based monitoring approach was formulated (Bolton and Hand, 2001; Byttner et al., 2011; Lapira, 2012), where the reference operational behavior of a particular substation in a network does not need to be predefined. Instead, its operational behavior is tracked by a local group of other similar substations within the network. In this sense, that particular substation is referred to as the target and the local group of other similar substations are referred to as its reference-group. Thus, the definition of a normal, atypical and faulty operational behavior in a target is now described relative to its reference-group. Under this setup, if the target is not behaving operationally in consort with the substations in its reference-group, then either it is due to a fault or because of some atypical operation arising at the target due to its local peculiarities. In effect, a reference-group acts as a just-in-time local model of a target's operational behavior. The reference-group based approach was demonstrated on the operational monitoring of the return temperature of 778 substations belonging to multi-dwelling buildings located in Helsingborg, Sweden. The results showed that this approach was able to detect deviations in the return temperature of substations over time by providing an adequate description of operational behavior for the target. That is, it provided for a comparison based on what the operational behavior of the target is and how its reference-group describes what it should be. Moreover, the approach was able to detect deviating targets which can be missed through setting a global threshold or the use of global models of outlier detection by providing for a more local context. Finally, based on the analysis, we created a categorization of most frequent deviation patterns observed in the return temperature data of the analyzed substations. This can be useful for creating a database of atypical and faulty operational behavior for this domain.

The remainder of this paper is organized as follows. Section 2 first gives an overview on the current state of the art of the data-driven analysis in the DH domain. Next, related work on the reference-group based monitoring approaches are discussed. Section 3 presents the reference-group based monitoring approach and summarizes it as an algorithm. Issues related to the reference-group, such as, its size and stability along with its adequacy to detect deviations at a local level are also addressed here. Section 4 provides a description about the data used in the case study. Details of setup for the study are then described. In Section 5, the results of the study are presented. Section 6 discusses the main findings of the study, including the limitations. Finally, the main conclusions of the study are presented in Section 7.

2. Related work

Various approaches have been proposed for the monitoring and fault detection of building related energy systems, for instance (Cai et al., 2014). For a comprehensive survey related to such approaches, the reader is referred to (Kim and Katipamula, 2018). However, the focus of this study is specifically on the application of group based monitoring to the DH domain. Therefore, the state of the art of data-driven analysis and fault detection for this domain is presented next. This is then followed by the state-of-the-art approaches in group based monitoring.

2.1. The state of the art of data-driven analysis and fault detection in DH

Most data-driven studies in DH have been on the prediction of entire network's heat demand. For instance, (Grosswindhager et al., 2011) proposes a model based on seasonal autoregressive integrated moving average (SARIMA) for short term on-line forecasting of the heat load in the DH network of Tannheim city, which is located in Tyrol in Austria. This article also studies outliers based on analyzing the model residuals by explicitly incorporating them into the model. Interestingly, this article also makes an observation about the challenges associated with the construction of models for individual substations:

Consumer load forecasting were not treated, due to the highly stochastic nature of the consumer data, which would make it necessary to build several individual models. Another such study in (Fang and Lahdelma, 2016) evaluates multiple linear regression models together with a SARIMA based model for forecasting the heat demand for the DH network of Espoo city in Finland.

Fewer studies have been done on substation meter data compared to electricity meter data as has also been observed in (Gianniou et al., 2018; Tureczek et al., 2019). The available literature is mostly concentrated on the clustering of substations based on their thermal energy use profiles. These clusters, in most studies, are then further analyzed for the purpose of fault detection. Interestingly, there are also not many studies related to heat-load forecast of individual substations. One such study in (Protić et al., 2015) proposes a multi-step heat-load forecast model for a substation based on a combination of support vector machine (SVM) and discrete wavelet transform (DWT). The data used in this study was obtained from one of the 3795 substations located in the DH network of Novi Sad, Serbia.

In a recent article (Månsson et al., 2019), interviews and surveys have been conducted on how DH utilities in Sweden perform fault detection on substations in their network. Additionally, statistics on commonly occurring faults as observed by DH utilities in various equipment of a substation have also been reported. It has been observed that analysis of the return temperature based on thresholds is the most commonly used control check for inefficient substation used by DH utilities. It has been further observed that other reported ways of detecting inefficient substations also mostly rely on threshold based methods. A study in (J. Pakanen and Ahonen, 1996) has already discussed the problems associated with a threshold based approach in DH way back in 1996. A later study in (Sandin et al., 2013) reiterated those results. Using statistical methods, the aforementioned study examines the operational behavior and conducts fault analysis on 996 substations located in Stockholm. To improve the operational efficiency of the entire DH networks, (Gadd and Werner, 2015) puts an emphasis on the need of understanding each individual substation in the network, with focus on fault detection. The aforementioned study performs fault analysis on 135 substations located in Helsingborg and Ängelholm, in Sweden. In particular, it identifies several types of faults and concludes that 3 out of 4 substations have some kind of faulty behavior. For improving fault detection, it also proposes to use thresholds for each substation based on the knowledge of customer's behavior. An approach based on partition around medoids (PAM) (Ma et al., 2017) uses heat load variation rather than heat load magnitude to group together similarly behaving buildings. The analysis is based on thermal energy use data collected from 19 higher education buildings located at Norwegian University of Science and Technology (NTNU) in Trondheim, Norway. One conclusion of the aforementioned study is that identifying daily heating energy use characteristics can be used to assist in fault detection and diagnosis strategies. Another study in (Xue et al., 2017) applies different clustering methods to extract heat load patterns and then uses association analysis to generate a set of association rules. It then uses these to detect faulty and energy inefficient substations. The data for the aforementioned study came from both the primary and the secondary side of two substations located in Changchun city, China. A k-means clustering approach is used in (Gianniou et al., 2018) to group and study the heat load profiles of 8293 single-family households in Aarhus, Denmark. This study finds that building's age and its area has a significant influence on its thermal energy use. The effect of autocorrelation on k-means clustering of the thermal energy use by heat exchange stations has been recently studied in (Tureczek et al., 2019). The data for this study came from 53 heat exchange stations located in Aarhus, Denmark.

Studies related to the distribution temperature of substations are also rather scarce. One such study in (Gadd and Werner, 2014) discusses different technical details about how to achieve a low return temperature. It analyzes 140 substations in Sweden, 85 located in Helsingborg and 55 in Ängelholm, to examine their temperature difference faults. The aforementioned study concludes that faults result in the increase of return temperature, which is then followed by an increase in the supply temperature. Another study in (Nord et al., 2018) analyzes technical possibilities for transitioning to low temperature district heating (LTDH). It proposes a thermal model with certain assumptions under which the supply temperature of the DH network, despite the presence of faults, can be reduced to 50°C and thereby also reducing the return temperature. This conclusion is based on the analysis of data from two DH networks located in Trondheim, Norway.

2.2. Group based monitoring and outlier detection

The idea of grouping similarly behaving systems in a large-scale deployment, though scarce, has been previously studied in (Bolton and Hand, 2001; Byttner et al., 2011; Das et al., 2010; Fan et al., 2015; Fontugne et al., 2013; Lapira, 2012; Narayanaswamy et al., 2014; Räsänen et al., 2008; Weston et al., 2012). A *peer-group* or reference-group based approach for detecting outliers among systems is proposed in (Bolton and Hand, 2001). The main idea here is to create a reference-group for each target object based on the k most similar objects criterion using Euclidean as a similarity measure. The application area of aforementioned study is credit card fraud detection and its assessment is done on the basis of visual inspection. Weekly spending data from 858 credit card accounts for a year is analyzed. Target accounts that moved further away from their respective peer-groups (or reference-groups) based on some externally defined threshold were considered suspicious. Using the same approach, a subsequent study in (Weston et al., 2012) examines faulty behavior in 200 weather stations. Another work in the area of fleet monitoring has been conducted by NASA for enhancing aviation safety (Das et al., 2010). The aforementioned work is based on detecting anomalies in multivariate flight operations quality assurance (FOQA) data. The study notes that the state-of-the-art approaches of the time were unable to deal with heterogeneity in the data containing both discrete and continuous variables. To overcome this limitation, the study proposes a multiple kernel learning (MKL) approach. Here, two kernels, one for discrete data and the other continuous data are combined linearly. It has been claimed that the MKL based approach is not only able to detect significant anomalies that were detected by the state-of-the-art approaches, but also other operationally significant anomalies in the data. Another study in (Fontugne et al., 2013) analyzes data of electrical energy use of devices in two different buildings with 135 and 70 room sensors. In each case, the data is divided into consecutive time bins and a pairwise correlation matrix between device's energy use for each time bin is computed. A constant reference matrix representing normal behavior is then created by computing the median of these correlation matrices. Faults are detected by applying a threshold to Minkowski weighted distance between the target and the reference correlation matrices at each time bin. Assessment of the faults is done on the basis of "known" fault signatures. True/False positive rates were not evaluated exhaustively due to lack of input from the domain expert. In another study (Narayanaswamy et al., 2014), parameters of energy use models for heating, ventilation, and air conditioning (HVAC) zones in a building were computed and then visualized using principal component analysis (PCA). The use of PCA revealed similar zones to be close to each other. Therefore, these PCA components were grouped together using k-means++ clustering. The resulting clusters were then examined by introducing three different definitions of fault. Overall, there were 237 HVAC zones, with 17 sensors each. Faults found in the study were assessed based on manual inspection of sensor data by a domain expert. Another study in (Lapira, 2012) uses clustering on 34 servo-gun units belonging to 30 industrial welding robots before the application of a fault detection procedure. Similar analysis is performed on 11 wind turbines from three wind farms. The aforementioned study (Lapira, 2012) shows that a clustering based approach to group similar behaving systems before the application of an outlier detection step is more robust to false alarms compared to individual models for each system. The so-called consensus self organized models (COSMO) (Byttner et al., 2011), a reference-group based monitoring approach, assumes the systems in a fleet to be operationally similar to each other. Here, each target system is compared to every other system in the fleet, i.e., its reference-group, and those found to be deviating the most based on a certain threshold are marked as outliers. This approach is later evaluated on a fleet of Volvo buses in (Fan et al., 2015) to detect those that behave differently from the group over time. A study in (Räsänen et al., 2008) identifies electricity usage patterns by using self organized maps (SOM) to create groups based on a list of building characteristics. The purpose, however, is geared towards studying the behavior of the building rather than fault detection. Nonetheless, the results of the study show the importance of meta-information about a system, which in most cases is not available. In a recent paper (Iyengar et al., 2018), building's physical attributes, e.g., its construction year and type, were used to create peer-groups for analyzing their energy efficiency. Those building groups with less than 20 houses were discarded on the pretext that this size is not enough for a meaningful analysis.

3. Methodology

Constructing a model for monitoring each system in a large fleet may not always be feasible. An alternative to this is fleet based monitoring. A fleet, according to (Oza and Das, 2012), is described as follows:

A fleet is a group of systems (e.g., cars, aircraft) that are designed and manufactured the same way and are intended to be used the same way.

The main assumption of a global model at the fleet level as described in (Oza and Das, 2012) is as follows:

Each system in the fleet is comparable to a sample drawn from some distribution, so that all the systems in the fleet are independent and identically distributed.

Thus, a global model assumes the operational behavior of all the systems in a fleet to be consistent with each other, i.e., homogeneous. Therefore, the behavior of any particular system can be inferred from the behavior of other systems in the fleet. Any system whose behavior is significantly different from the fleet is considered as an outlier. This is referred to as group based or fleet based monitoring (Byttner et al., 2011; Oza and Das, 2012). In this context, the entire fleet or network of systems is the reference-group. However, in many situations, there can be differences in the ambient environment, installation and control settings, model type, age and many other factors among the systems in a fleet, details about which may not be readily available. Under these conditions, a global model based on the homogeneity assumption will not be very efficient since it leads to the detection of only those systems that are outliers in the global sense. Therefore, the notion of similarity may need to be relaxed here as described in (Bolton and Hand, 2001; Lapira, 2012), which specifies that similarity among systems does not necessarily imply that they are exactly identical. Hence, a system in a fleet should be compared to only those systems that are found to be the most similar to it in terms of their behavior. This context of a peer-group or a reference-group based approach has been described in (Bolton and Hand, 2001):

The distinguishing feature of peer-group analysis lies in its focus of local pattern analysis rather than global models.

In addition, a motivation for the reference-group based approach also comes from quantitative biochemistry (Livak and Schmittgen, 2001), according to which:

Relative quantification describes the change in expression of the target gene relative to some reference group such as an untreated control or a sample at time zero in a time-course study.

3.1. A reference-group based approach for detecting outlier systems

In the context of monitoring a large fleet, the main task is to select an appropriate method to construct a reference-group for each target system. The issue that arises here is that in most cases, the behavior of systems in such a fleet lies on a spectrum and not necessarily consists of some discrete sets of well separable values. This makes it difficult to apply a clustering based approach. Moreover, as noted earlier, meta-information about those systems that can be useful to compare or distinguish them from each other is not always available. In addition, clustering based methods such as the k-means and the Gaussian mixture model (GMM) impose a certain distributional criterion on the underlying data distribution of the features selected to represent the systems in a fleet, which may not reflect the ground truth. A k-nearest neighbor (k-NN) based criterion does not impose such a restriction. From the point of view of outlier detection, it has been shown in (Goldstein and Uchida, 2016) that an imperfect choice of k tends to give more stable results for a nearest neighbor based approach than a clustering based approach. Moreover, the basic principle of nearest neighbor as pointed out in (Cover, 1982) is: *things that look alike must be alike*, which is a requirement for a reference-group. Due to these aforementioned factors, the notion of k-nearest systems based on the definition of an appropriate similarity measure is justified.

Consider a large fleet (or network) of N systems and let $z_{i,1}$ represent the (state of) the i-th system at time $t = 1$. The evolution of the fleet's behavior over time can be represented by a matrix:

$$\zeta_{N\times T} = \left[\underbrace{\begin{matrix} z_{1,1} & z_{1,2} & \cdots & z_{1,t} \\ \vdots & \vdots & \ddots & \vdots \\ z_{N,1} & z_{N,2} & \cdots & z_{N,t} \end{matrix}}_{\substack{\text{reference-group creation} \\ \text{time-period: } \xi_0=\{1,...,t\}}} \; \underbrace{\begin{matrix} z_{1,t+1} & \cdots & z_{1,t+s} & \cdots & z_{1,T} \\ \ddots & \vdots & \ddots & \vdots & \\ z_{N,t+1} & \ddots & z_{i,t+s} & \cdots & z_{N,T} \end{matrix}}_{\substack{\text{reference-group evaluation} \\ \text{time-period: } \xi_1=\{t+1,...,T\}}} \right] \tag{1}$$

Any new data that arrives at a latter time, i.e., $T + 1$, $T + 2$, and so on, can be appended to the above matrix. For simplicity, assume that each system is represented by a single variable. Moreover, let $\xi_{m=0} = [1, ..., t]$ represent the 0-th episode. Then,

$$\mathbf{y}_{i,\xi_0} = [z_{i,1}, ..., z_{i,t}] \qquad i = 1, ..., N. \tag{2}$$

Each particular system in the fleet which is selected for the purpose of its operational monitoring is referred to as a target. In general, systems in a fleet may differ due to factors, such as, their ambient environment, control settings, etc. This information, as discussed earlier, may not always be available. However, assume that there exist sufficient number of systems, which share a similar but unknown underlying data generating mechanism. Further, assume that there exists an appropriate distance measure that can adequately estimate their similarity. Then based on the definition of the selected distance measure, the reference-group comprises of those systems whose underlying data generating mechanism is approximately similar to that of the target. Assume that the data from episode $\xi_{m=0}$ is sufficient to capture the similarity among the systems in a fleet, and that the Euclidean is an appropriate distance measure for this task. Then, the similarity between systems i and j can be measured as follows:

$$d(i, j) = \sqrt{(\mathbf{y}_{i,\xi_0} - \mathbf{y}_{j,\xi_0})(\mathbf{y}_{i,\xi_0} - \mathbf{y}_{j,\xi_0})^*}. \tag{3}$$

The application of $d(\cdot,\cdot)$ to estimate the similarity between N systems results in a $D_{N\times N}$ distance matrix consisting of pairwise distances. Each row i of D is then sorted in the decreasing order of similarity (or increasing distance). Let some $k << N$ be the appropriate size of the reference-group. Then, the indexes of the reference-group for each target system i sorted in the decreasing order of similarity are given by:

$$\pi^i_{1:k+1} = \text{argsort } D_{i,:} \qquad i = 1, ..., N, \tag{4}$$

where $\pi^i_1 = i$ is the index of the target in matrix D. These index vectors are created at the 0-th episode, i.e., ξ_0, and then kept fixed. Let $\tau_i = \mathbf{y}_{\pi^i_1,\xi_0}$ represent the target and $\mathbf{r}_i = [\mathbf{y}_{\pi^i_2,\xi_0}, ..., \mathbf{y}_{\pi^i_{k+1},\xi_0}]$ its reference-group. The main assumption here is that over subsequent episodes ξ_1, ξ_2,..., the same set of unobserved variables, say Υ, will continue to affect the target i along with its reference-group. Therefore, the target i will continue to behave in a similar manner compared to most of the systems in its reference-group. Then under certain algorithmic criteria Θ, if the target i is found to be behaving differently from its reference-group, it is considered an outlier.

The nature and quality of outlier detection will depend of various factors including the choice of the similarity measure used to create the reference-groups, the notion of stability of the reference-group, the size k of the reference-group, the adequacy of the reference-group to discover outlier targets at the local level, the representation of the target and its reference-group along with choice of outlier detection procedure Θ. We discus each one of these in the following sections.

3.2. The choice of the similarity measure, stability and the size k of the reference-group

The first step in the creation of a reference-group is the choice of a similarity measure. This selection can be made on the notion of stability based on the following criterion: The stable proportion of a particular distance measure for a given k is the ratio between the systems that were members of the reference-group at the episode of its creation that continue to remain its members when it is reconstructed in the next subsequent episode, and k. An adequate threshold, say $\delta_s \in [0,1]$, can be used to select the desired stable proportion. The stable proportion of a reference-group is usually an increasing function of its size k. The similarity measure which fulfills this threshold criterion with the least reference-group size k is selected. Hence, using the stability criterion, the size k of the reference-group can be simultaneously determined. Hence, for each target i, the size of its reference-group is given by $k_i = k'$ when the following approximation holds:

$$\frac{\#\left[\pi^i_{2:k'+1,\xi_0} \bigcap \pi^i_{2:k'+1,\xi_1}\right]}{k'} \approx \delta_s \quad k' = 1, ..., N, \qquad i = 1, ..., N. \tag{5}$$

Here, # is the cardinality of the set and $\bigcap$ is the set intersection operator. Based on the formulation of Eq. (5), a reference-group for each target i will have a different size k_i. A single global k, though not efficient, can be selected by taking the median of the vector given by:

$$k = median\ [k_i]^N_{i=1}. \tag{6}$$

The loss of members in the reference-group over the next episode is usually driven by the changes in the data distribution of its members. A desirable reference-group should be stable over time.

Concerning the question of minimum k, (Breunig et al., 2000) proposes it to be at least 10 to avoid statistical fluctuations. Yet another perspective, described in the same paper is based on the definition of a local cluster. It has been suggested that this cluster must contain at least k objects (the reference-group in our case) so that other objects (the target, in our case) can be outliers relative to the cluster. In this sense, the choice of minimum k is application specific. In summary, each member of a reference-group adds to the evidence about a particular behavior that its target is supposed to follow.

3.3. Adequacy of a reference-group

An adequate reference-group is one which follows its target as closely as possible. In this regard, an adequacy measure, which estimates the similarity between the target and its reference-group can be useful. Such a measure can be based on some distributional distance or correlation. For instance, the median distributional distance between the target and each member of its reference-group can be estimated by:

$$\bar{H}_i = median\ [\Omega(\mathbf{y}_{i,\xi_m}, \mathbf{y}_{\pi^i_j,\xi_m})]_{j=2}^{N}, \qquad i = 1, ..., N. \tag{7}$$

If the underlying data of each member in the reference-group and the target is assumed to follow a Gaussian distribution, Ω can be described by Hellinger distance via its relation to Bhattacharyya coefficient (BC):

$$\Omega(p, q) = \sqrt{1 - BC(p, q)}, \quad p \sim N(\mu_1, \sigma_1^2), q \sim N(\mu_2, \sigma_2^2), \tag{8}$$

where $BC = \sqrt{\frac{2\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2}} \exp^{-\frac{1}{4}\frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2}}$. In case the data distribution of each member in the reference-group and the target is not known, Ω can be described by the Kolmogorov-Smirnov (KS) two-sample test, used as a distance measure. KS and Hellinger distance based on BC are bounded in the range [0, 1]. KS measures the maximum distance between the cumulative density functions of the two distributions. Hellinger distance via its relation to BC measures the amount of non-overlap between the two distributions. In this sense, an adequate reference-group is one where the median distributional distance between its members to the target is less some specified threshold.

Let the mean behavior of the members of the reference-group be represented by:

$$\mathbf{u}_{i,\xi_m} = \frac{1}{k}\sum_{j=2}^{k+1} \mathbf{y}_{\pi^i_j,\xi_m} \qquad i = 1, ..., N. \tag{9}$$

Then the correlation between the mean behavior of the reference-group and its target can be estimated by:

$$\bar{\rho}_i = \rho(\mathbf{u}_{i,\xi_m}, \mathbf{y}_{i,\xi_m}), \qquad i = 1, ..., N. \tag{10}$$

Accordingly, an adequate reference-group here is one that strongly follows in the direction of its target over time, based on some specified threshold.

Let $\delta_{\bar{H}}$ and $\delta_{\bar{\rho}}$ be the minimum acceptable distance and correlation thresholds, respectively. These are referred to as the adequacy criteria. When $\bar{H} > \delta_{\bar{H}}$, it implies that the underlying parameters such as the mean or standard deviation of the target and/or the majority of the members of its reference-group have changed. Similarly, when $\bar{\rho} < \delta_{\bar{\rho}}$, it implies the target and its reference-group do not behave in a similar fashion over time.

When $\bar{\rho}$ and $\bar{H}$ violate their respective thresholds at the episode of the creation of a reference-group, it might indicate that the target already has a fault or it has a unique behavior which is not mirrored by other systems. In the latter case, a reference-group based approach may not be applicable for the particular target. When a violation occurs in the next subsequent episodes, it might be due to a fault or due to changes in the target, such as changes in its working environment or control setting. In the latter case, a new reference-group may be required for the target.

3.4. Representing the target and its reference-group

Once an appropriate similarity measure has been chosen and the reference-group for each target identified, appropriate features can be computed to represent the target and its reference-group. For instance, the mean of the target i along with the mean of each member of its reference-group over some episode $\xi_m = [1, ..., W]$ is given by:

$$u_j^{\mu} = \frac{1}{W}\sum_{\xi_m=1}^{W} \mathbf{y}_{\pi^i_j,\xi_m}, \qquad j = 1, ..., k+1, \tag{11}$$

where u_1^{μ} is the mean of the target. Similarly, the standard deviation is given by:

$$u_j^{\sigma} = \sqrt{\frac{1}{W-1}\sum_{\xi_m=1}^{W} (\mathbf{y}_{\pi^i_j,\xi_m} - u_j^{\mu})(\mathbf{y}_{\pi^i_j,\xi_m} - u_j^{\mu})^*}, \qquad j = 1, ..., k+1, \tag{12}$$

where u_1^{σ} is the standard deviation of the target. The target i along with its reference-group can then be represented by:

$$\tau_i = [u_1^{\mu},\ u_1^{\sigma}] \qquad \mathbf{r}_i = [u_j^{\mu},\ u_j^{\sigma}]_{j=2}^{k+1} \qquad i = 1, ..., N. \tag{13}$$

Other features, such as min, max, skewness, kurtosis can also be added to Eq. (13).

3.5. The outlier detection procedure Θ

The choice of an outlier detection procedure Θ can be an unsupervised outlier detection method, such as, the isolation forest (IF) (Liu et al., 2012), using for instance, Eq. (13), as an input. Other unsupervised outlier detection methods such as one class-support vector machine (OC-SVM) (Schölkopf et al., 1999) can also be used here assuming that the reference-group represents the normal behavior.

In this study, IF is chosen as the outlier detection procedure Θ. This choice is based on the study conducted in (Emmott et al., 2015), which after comparing eight outlier detection methods reaches the following conclusion: *Because Isolation Forest performed best on average and because it has very good runtime properties, we recommend it for general use. However, we also recommend that context should impact your choice of algorithm.* In addition, a recent study in (Domingues et al., 2018) after comparing fourteen outlier detection methods also makes a similar conclusion.

IF works on the assumption that outliers are easily isolated compared to normal data points. In the context of this study, the requirement is to check if the target is isolated among its reference-group. In certain instances, a reference-group may itself contain outliers. To deal with the issue, the contamination rate of IF can be adjusted. So, if the reference-group consists of 40 systems, then for a contamination rate of 0.1 or 10%, it can be checked if the target is among the list of those 4 systems that IF considered as isolated. If this happens to be the case, then the target is an outlier.

On application of $\Theta(\tau_i, \mathbf{r}_i)$, if the target is found to be behaving differently from its reference-group, it is considered an outlier. When the adequacy criteria discussed earlier are not met, Θ in many cases can fail to detect an outlier target. In cases where a target is found to be an outlier by Θ, the adequacy criteria can be used in conjunction to understand possible reasons behind its outlierness.

The steps required for detecting outlier targets using a reference-group based approach are summarized in Algorithm 1.

Algorithm 1 Reference-group based deviation and outlier detection

Require: Time-series data $[\mathbf{y}_{1,\xi_m}, ...\mathbf{y}_{N,\xi_m}]$ from N systems corresponding to episodes $[\xi_0, \xi_1, ...]$, an appropriate distance measure $d(\cdot,\cdot)$ to estimate similarity between systems, the size k of the reference-group, a function Γ that computes an appropriate representation for the target and its reference-group, an outlier detection procedure Θ that outputs -1 if the system is an outlier and 0 otherwise.

```
for i ← 1 to N do
    for j ← 1 to N do
        D_{i,j} = D_{j,i} = d(y_{i,ξ0}, y_{j,ξ0})
    end for
    π^i_{1:k+1} = argsort D_{i,:}  //the first item of π is the index of target system.
end for
for m ← 0 to ∞ do
    for i ← 1 to N do
        τ_i, r_i = Γ(y_{π^i_{1:k+1}, ξm})
        β_i = Θ(τ_i, r_i)
        if β_i = −1 then
            The target system i is an outlier.
        end if
    end for
end for
```

4. Case study

The focus of this study was the return temperature of substations since it affects efficiency of the entire DH network. Historically, in Sweden, the most commonly used variable for analyzing substations used to be the temperature difference, i.e., the difference between the supply are the return temperature. This was because in earlier times, only heating rate and flow rate data were collected, from which the temperature difference was indirectly estimated. However, since the supply temperature is not constant over the year and differs between networks, the temperature difference may not be a reliable variable, especially when comparing substations with each other. Therefore, it is much better to use return temperature, since there is a physical lower limit defined by the indoor temperature.

The choice of multi-dwelling buildings for the analysis rests on the fact that according to Öresundskraft, it represents more than 50% of the heat deliveries (55% in Helsingborg) in their DH network. Moreover, the overall market share of DH in this category is 93%.

4.1. Data description and preprocessing

The dataset used in this study was provided by Örsundskraft, a DH utility located in the South-West of Sweden. It was derived from smart meter readings from buildings connected to the DH network of the Helsingborg municipality. In 2017, 3070 TJ of heat was delivered to 11,242 delivery points. The dataset included hourly measurements of the heating rate, the flow rate, the supply temperature and the return temperature on the primary side of all substations during 2017. Moreover, it was divided into six customer categories: single family buildings, multi-dwelling buildings, public administration buildings, commercial buildings, health and social service buildings and others.

The return temperature data of the multi-dwelling buildings from Nov'17 and Dec'17 was used in the analysis. In this respect, data from a total of 965 substations in Nov'17 and 963 substations in Dec'17 was available. For data preprocessing, substations with less than 85% hourly return temperature data points in each month were removed. Moreover, substations with constant values were also removed. For the rest, missing values, if any, were imputed using linear interpolation. Furthermore, only those substations having data from both the months, i.e., Nov'17 and Dec'17, were retained. The aforementioned data preprocessing left a total of 865 substations from the two aforementioned months.

The return temperature is a very volatile operational variable (Sandin et al., 2013). Therefore, to remove the influence of outliers and other variations, daily mean of the return temperature for each of the 865 substations was calculated before any further analysis.

4.2. Setup for the analysis

According to (Frederiksen and Werner, 2013), the national average return temperature in Sweden is 40°C-50°C. Therefore, a global threshold of 50°C can be considered as the red line that differentiates a sufficiently normal substation from a bad one. Hence, it was used as a global threshold in our analysis.

Five different distance based similarity measures were studied for creating a reference-group for each target substation: Euclidean, Wasserstein (Earth Movers Distance), Energy (Cramér von Mises), Hellinger and KS. The stability of the reference-groups based on the aforementioned distance measures for various values of k were tested using Eq. (5). A median size k was calculated using Eq. (6) and its selection was based on $\delta_s = 0.60$. Finally, the adequacy of the reference-groups were based on the following criteria: (1) The correlation between target and the mean behavior of members of its reference-group was $\bar{\rho} \geq 0.60$. (2) The median Hellinger distance between the target and members of its reference-group was $\bar{H} \leq 0.40$.

The month of November is usually the start of winter season in Sweden. It was assumed that the pairwise distances between substation's return temperature data from this month provides sufficient information on their similarities and diversities. Hence, the reference-group for each target was created using the return temperature data from Nov'17, i.e., episode ξ_0. The reference-groups were then kept fixed on the assumption that they will continue to follow their respective targets in a similar manner over the next subsequent episode. Therefore, the operational behavior of the return temperature for each target was observed relative to its reference-group for Dec'17, i.e., episode ξ_1. Monthly mean and standard deviation were used to represent the target along with its reference-group. This representation from Dec'17 was then used as an input to the IF method to detect the change in behavior of the target relative to its reference-group. The contamination rate of the IF was set to 0.10 or 10%. Hence, a target was considered an outlier if it was found to be among those 10% substations identified as such by IF. The overall process for detecting outlier targets using the reference-group approach was based on Algorithm 1.

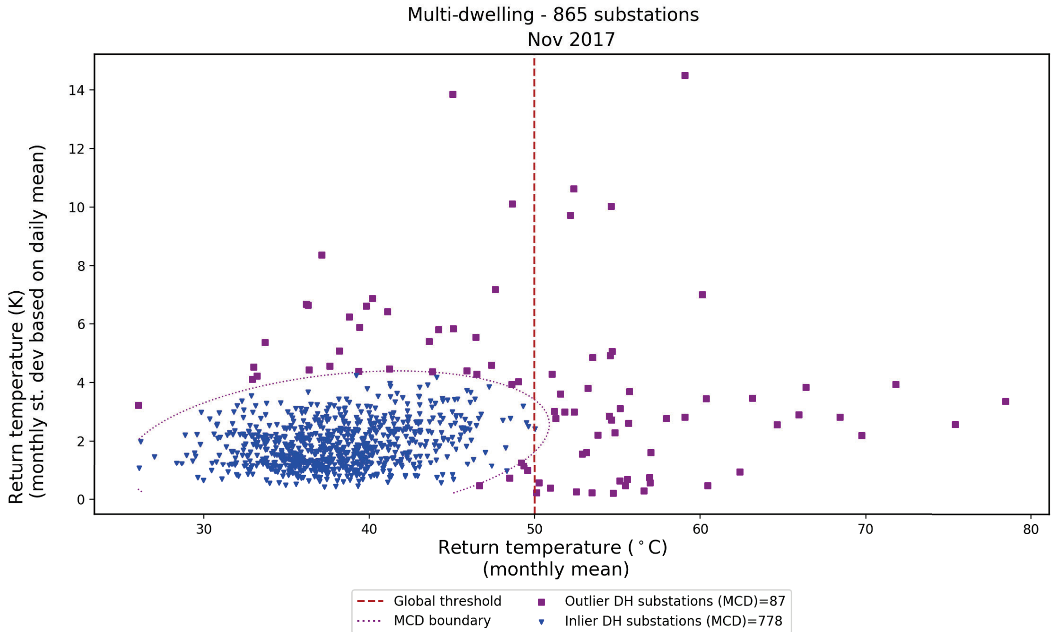

Figure 1: Based on the global threshold, 52 (5.9% contamination rate) substations had a monthly mean return temperature greater than 50°C in Nov'17. MCD discovered 87 outlier substations based on a contamination rate of 0.1 or 10%.

The programing environment used in this study was Python 3.7. All distance measures, except for Hellinger, were computed using the scipy library (ver. 1.3.0). The Hellinger distance was computed using Eq. (8). The outlier detection methods based on IF and minimum covariance determinant (MCD) (Rousseeuw and Van Driessen, 1999) were used from the scikit-learn library (ver. 0.20.0). MCD is a distance based statistical outlier detection method, which was used to create a global outlier model for all substations.

5. Results

Figure 1 shows the monthly mean and standard deviation of the return temperature of substations based on Nov'17 data. As noted before, the state of practice in DH industry is to set global thresholds. A normally accepted threshold for the return temperature is around 50°C. As can be observed in Figure 1, the dashed dark red line separated those substations which had a return temperature higher than 50°C from the normal ones. In total, there were 52 substations with a return temperature higher than 50°C.

High standard deviation in the return temperature can be another sign of problem in a substation. A global outlier model based on the MCD was applied to the representation shown in Figure 1. Based on the criterion that the decision boundary of the MCD model agreed with the global threshold, the contamination rate for MCD was set to 0.1 or 10% (top outliers). With that, 87 substations with the return temperature of around 50°C or higher and with a standard deviation of around 4 K (Kelvin) or higher were identified as global outliers. These substations were removed from both Nov'17 and Dec'17, leaving the total to 778. Individual analysis of some of these substations by the DH expert suggests that those with return temperature of less than 50°C but with a standard deviation greater than 5 K were not necessarily faulty. In particular, some of these substations switched on a time-clock operation for ventilation (Frederiksen and Werner, 2013), which also affected the return temperature by increasing their standard deviation.

5.1. Stability analysis of reference-groups

Reference-groups were created for the 778 target substations using different distance measures and various values of k for episode ξ_0 and ξ_1. Figure 2 shows the stable proportion of the reference-group for a given distance measure and median k between episodes ξ_0 and ξ_1. The Euclidean distance reached the threshold of $\delta_s = 0.6$ with the minimum median k, i.e., 80, compared to other distance measures. Moreover, for the Euclidean distance, increasing median k

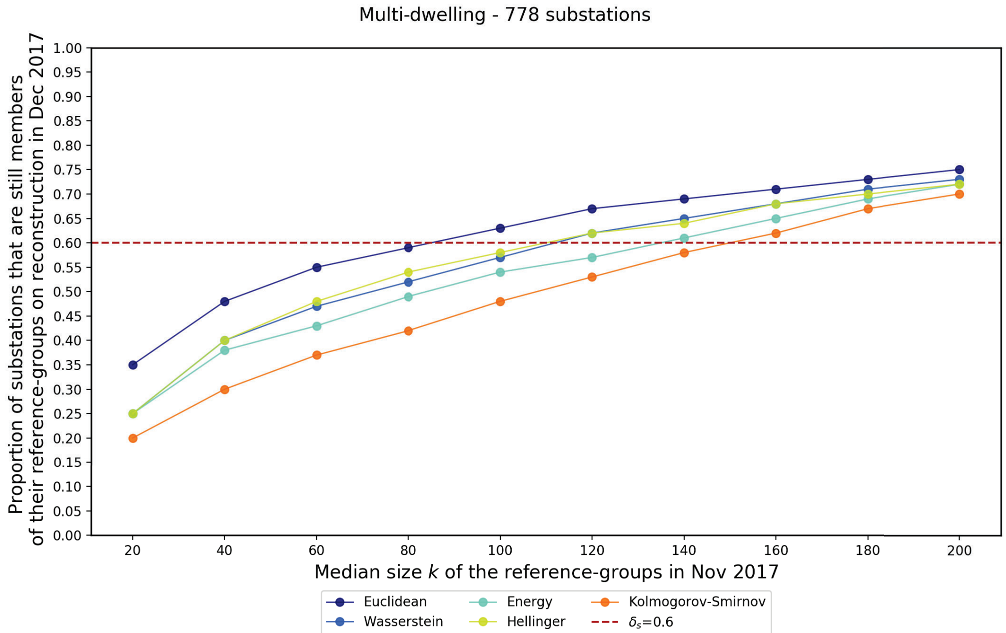

Figure 2: On the x-axis is the median size k of the reference-groups at the time of their creation. On the y-axis, is the proportion of reference-group members that if it was reconstructed in the next episode, were still its members. Based on this criterion, the Euclidean distance appeared to be more stable among all other distance measures.

beyond 80 did not show large marginal increases in the stable proportion. Hence, Euclidean distance with $k = 80$ appeared to be a reasonable choice for creating the reference-groups.

5.2. Adequacy analysis of reference-groups

Once the reference-groups for their respective targets were created in episode ξ_0 i.e., Nov'17, they were kept fixed for the next episode ξ_1, i.e., Dec'17. However, dispersion among the members of the reference-groups over time can reduce their effectiveness. This is obvious, because there are many unobservable factors that can change at the substations or the particular buildings they are associated with. For instance, some substations might have changed their outdoor temperature compensation curve, people may have moved into/out of the buildings, some sort of equipment fault might have occurred at substations, additional heaters were removed/introduced in the buildings, etc.

Figure 3 shows the histograms of the adequacy measures for the reference-groups constructed using Euclidean distance with $k = 80$. It shows that on aggregate basis, median $\bar{\rho}$ decreased from 0.85 to 0.81, while median $\bar{H}$ increased from 0.21 to 0.30, from Nov'17 to Dec'17. Reference-groups not fulfilling $\bar{\rho} \geq 0.60$ showed an increase from estimated 175 (22%) in Nov'17 to 210 (26%) in Dec'17. Similarly, reference-groups not fulfilling $\bar{H} \leq 0.40$ showed an increase from 68 (8%) in Nov'17 to 201 (25%) in Dec'17. Possible reasons for the reference-groups not fulfilling these adequacy criteria include: (a) the reference-groups could not adequately represent their respective targets, and (b) the targets behaved differently from their reference-groups and therefore might have developed a fault. In the former case, an outlier detection model may miss to detect a deviating target.

5.3. Detection of outlier target substations

The results of Table 1 were obtained using the reference-group based approach described in Algorithm 1. Moreover, although Euclidean distance with $k = 80$ was selected for the creation of the reference-groups based on the analysis in Figure 2, the results of different distance measures for various values of the size k of the reference-groups can provide some additional useful insights. Against each distance measure in Table 1, the first row consists of global outlier targets, the second row consists of local outlier targets and the third row consists of total outlier targets, i.e., it is the sum of global and local outlier targets. The fourth row consists of global outlier targets detected by the MCD model but missed by the reference-group based approach. On an overall basis, it can be observed that the total number of outlier targets

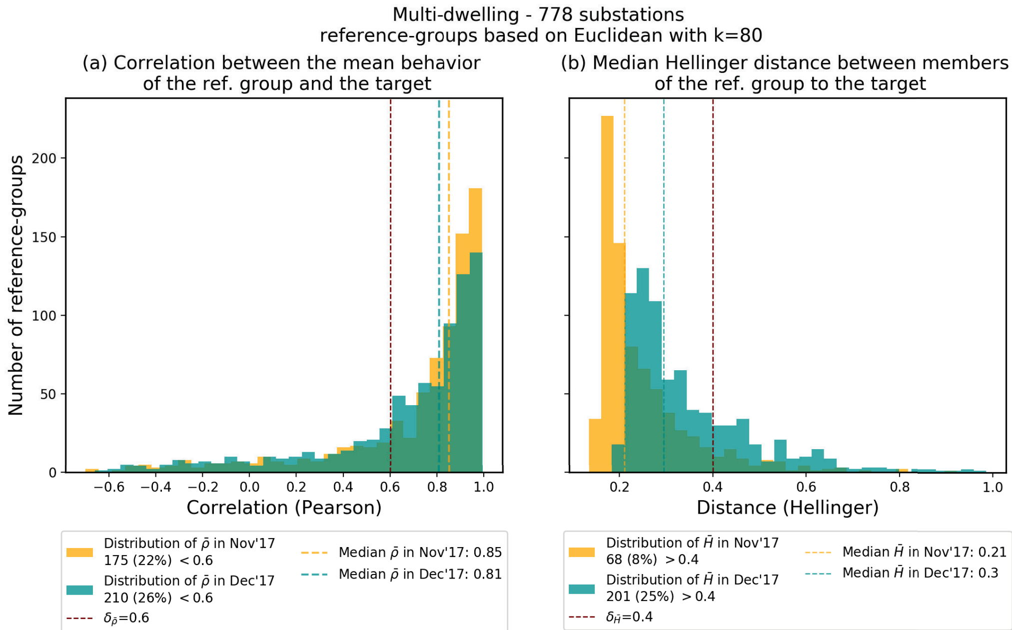

Figure 3: (a): Here, the histograms based on correlations between the target substations and the mean behavior of their respective reference-groups are shown. When the reference-groups were created at Nov'17, 175 (22%) of them did not fulfill the adequacy criterion based on correlation. This increased to 210 (26%) in Dec'17. (b): Here, the histograms based on median Hellinger distances between each member of the reference-groups to their target are shown. About 68 (8%) of the reference-groups did not meet the adequacy criterion based on Hellinger distance in Nov'17. This increased to 201 (25%) in Dec'17.

detected may depend both on the choice of the distance measure as well as the size k of the reference-group. Moreover, it can also be observed that at lower values of k, more local outlier targets were detected compared to the global outlier targets. However, this also increases the possibility of false alarms. With increasing k, the relative proportion of the local outlier targets decreases, while that of the global outlier targets, it increases. Hence, a trade-off appeared to exist between the number of global and local outlier targets depending on the size k of the reference-group. Finally, with increasing k, less global outlier targets were missed compared to the global MCD model. Hence, as expected, with increasing k, the reference-group approach moved towards becoming a global model.

5.4. A comparison between reference-group based and global outlier detection approaches

Figure 4 presents a comparison of the reference-group based approach with global threshold based and MCD based global outlier models. Using the criterion that the boundary of the MCD model agreed with the global threshold of 50°C, the contamination rate for MCD for Dec'17 was set to 5%. The MCD model detected 39 global outlier targets including five out of six detected by the global threshold.

The results for the reference-group based approach in Figure 4 were obtained using the Euclidean distance with $k = 80$. In comparison to global models, the reference-group based approach detected 33 global outlier targets, missing six of them. However, as can be observed in Figure 4, those missed were close to the boundaries of MCD and the global threshold. In addition, the reference-group based approach detected 44 local outlier targets which can be observed as red dots inside the MCD boundary in Figure 4. Although these additional outliers increased the overall contamination rate to 10%, they presented the possibility of detecting potentially problematic cases at a local level, albeit, at the cost of false alarms. In this regard, a 5% increase in the overall contamination rate is tolerable.

To illustrate further on these local outlier targets, an example is presented in Figure 5. Observe the target, marked as a red dot, along with its reference-group marked as light blue dots. The target together with its reference-group consisting of 80 substations had an average return temperature of 35°C in Nov'17. However, in Dec'17, the target showed a return temperature of 40°C, which was a significant increase compared to the majority of those in its reference-group.

Distance		Outliers									
		k=20	k=40	k=60	k=80	k=100	k=120	k=140	k=160	k=180	k=200
Euclidean	Global	30	31	33	33	35	34	37	37	36	37
	Local	45	52	51	44	48	45	43	45	37	35
	Total	75	83	84	77	83	79	80	82	73	72
	Missed	9	8	6	6	4	5	2	2	3	2
Wasserstein	Global	25	25	27	32	33	34	34	36	35	33
	Local	52	47	45	42	35	35	39	37	32	34
	Total	77	72	72	74	68	69	73	73	67	67
	Missed	14	14	12	7	6	5	5	3	4	6
Energy	Global	26	25	28	30	33	33	33	31	37	34
	Local	47	43	37	32	34	34	37	37	34	36
	Total	73	68	65	62	67	67	70	68	71	70
	Missed	13	14	11	9	6	6	6	8	2	5
Hellinger	Global	21	24	25	25	30	29	28	31	32	32
	Local	45	38	35	36	39	34	33	32	33	35
	Total	66	62	60	61	69	63	61	63	65	67
	Missed	18	15	14	14	9	10	11	8	7	7
KS	Global	22	26	25	28	31	32	30	32	33	33
	Local	48	37	35	39	40	32	37	37	35	34
	Total	70	63	60	67	71	64	67	69	68	67
	Missed	17	13	14	11	8	7	9	7	6	6

Table 1
Outliers detected in Dec'17 by tracking the 778 target substations with their respective reference-groups created using different distance measures. For each distance measure, the first row consists of global outlier targets, the second row consists of local outlier targets and the third row is the sum of global and local outlier targets. The fourth row consists of global outlier targets detected by the MCD model, but missed by the reference-group based approach.

With such an increase, there was a fair chance that there might be a problem. Even if there was no fault, such increase could be due to inefficient control setting at the substation. Hence, a potential problem was identified before it could be possibly detected by a global outlier detection approach at some later point in time. In many instances, if the situation is not dealt with in time, such cases have the potential of further deterioration where eventually the 50°C line is crossed. Other than that, it can also be observed in Figure 5, that a reference-group provided for a relative comparative reference to judge a target on its outlierness.

5.5. Return temperature patterns of outlier substations

Following the reference-group based approach, the observed pattern of behavior in the return temperature of the 77 outlier target substations can be summarized into the following:

1. Constant (5)
2. Fluctuating (9)
3. Temporary increase (15)
4. Temporary decrease (10)
5. Level increase (22)
6. Level decrease (9)

The observed occurrence of each pattern is stated in the brackets. For seven out of 77 cases, none of the above patterns seemed to fit clearly. The analysis of example cases associated with the six patterns above using the reference-group approach based on Algorithm 1 are presented next. Here, Euclidean distance with $k = 80$ was used for creating the

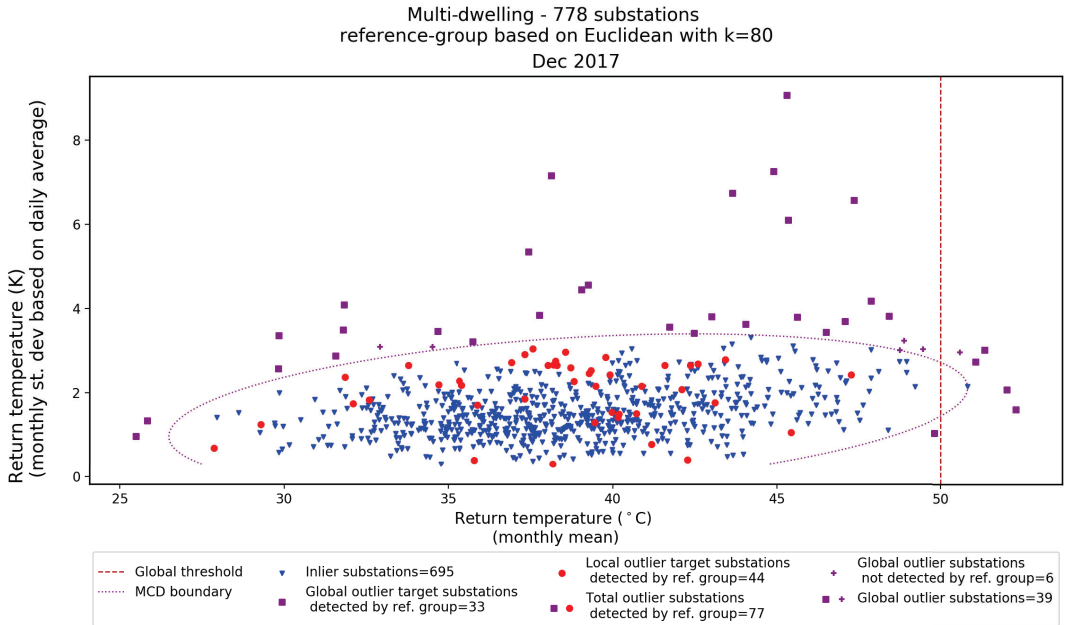

Figure 4: The reference-group based approach not only detected global outlier targets, but also those at the local level. It can also be observed that those missed global outlier targets were very close to boundaries associated with the global models.

reference-groups. However, the specific cases that are discussed were also found by all the other four distance measures with $k = 80$, except for the one associated with the "fluctuating" pattern: it was not detected with the Hellinger distance. We therefore considered that Euclidean with $k = 80$ was sufficient to describe the aforementioned six patterns with the example cases.

In Figure 6 to Figure 11, red color lines and dots represent the target substation, light blue color lines and dots represent the members of its reference-group and dark blue color lines and dots represent the mean behavior of the reference-groups. The dashed dark red line represents the global threshold at 50°C. The adequacy criteria based on correlation and Hellinger distance are also shown on each of these figures. The adequacy criteria can be useful in understanding some reasoning behind a target's outlierness.

Since no labels or ground truth on normal, atypical or faulty behavior were available, each example case associated with the six patterns was specifically discussed with the DH expert at Öresundskraft, Sweden. No additional information other than what can be observed in Figure 6 to Figure 11 was part of the discussion. In that way, the mathematical and practical knowledge were merged into the following description:

5.5.1. Example case of the "constant" pattern

Figure 6(a) shows that the return temperature of the target substation had an almost constant level with very low standard deviation compared to its reference-group during Nov'17. Moreover, a low $\bar{\rho}$ indicated that the reference-group did not follow the target well over time. Furthermore, $\bar{H}$ was at the threshold boundary of 0.4. All this indicated that the target could not be sufficiently represented by a reference-group. In Figure 6(b), it can be observed that the situation did not change much in Dec'17. According to the DH expert, even though the target substation was an outlier here, it does not appear to be a fault.

5.5.2. Example case of the "fluctuating" pattern

In Figure 7(a), both $\bar{\rho}$ and $\bar{H}$ indicated that the target could not be well represented by its reference-group. Moreover, the target showed erratic behavior compared to its reference-group. According to the DH expert, the behavior could be due to a malfunction or inefficient control setting at the target substation. In Figure 7(b), it can be observed that there was no improvement of situation in Dec'17.

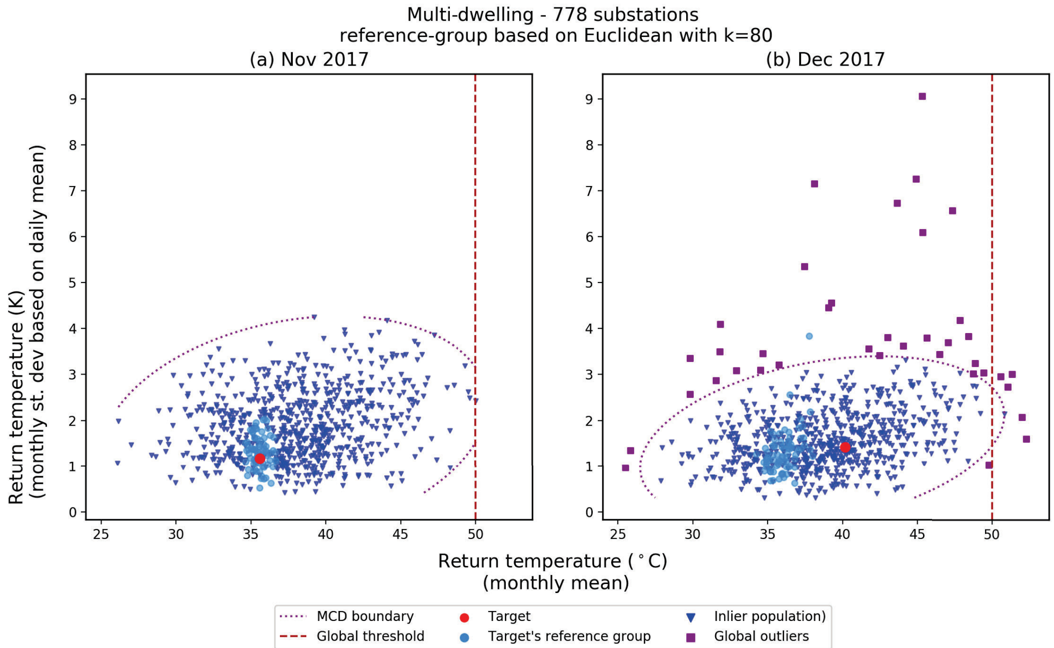

Figure 5: (a): The target substation (red dot) in Nov'17, relative to its reference (light blue dots) and population (blue dots) is shown. (b): The target substation relative its reference-group and population in Dec'17 is shown.

5.5.3. Example case of the "temporary increase" pattern

Here, in Figure 8(a), both $\bar{\rho}$ and $\bar{H}$ indicated that the target was well represented by its reference-group. In Figure 8(b), the target showed a sudden continuous increase in its return temperature for around six days before it reverted back to the behavior of its reference-group. Both $\bar{\rho}$ and $\bar{H}$, albeit near their threshold boundary, indicated that the target had deviated from its reference-group. According to the DH expert, a possible reason could be a control malfunction of a short duration at the target substation in Dec'17.

5.5.4. Example case of the "temporary decrease" pattern

In Figure 9(a), both $\bar{\rho}$ and $\bar{H}$ indicated that the target was well represented by its reference-group. In Figure 9(b), the target showed a decrease in its return temperature for a few days around the mid period of Dec'17, before later appearing to move closer to the behavior of its reference-group. Both $\bar{\rho}$ and $\bar{H}$, during this period indicated that the target had deviated from its reference-group. According to the DH expert, a possible explanation could be a significant change in heating use during the period. Other possible reasons could be a temporary change of control settings at the target substation.

5.5.5. Example case of the "level increase" pattern

In Figure 10(a), both $\bar{\rho}$ and $\bar{H}$ indicated that the target was well represented by its reference-group. Figure 10(b) shows that there was a sharp increase in the return temperature of the target compared to that of its reference-group in Dec'17. This resulted in an obvious deterioration of both $\bar{\rho}$ and $\bar{H}$ during the period. This behavior, according the DH expert, can possibly be due to a fault in the flow control valve at the target substation.

5.5.6. Example case of the "level decrease" pattern

In Figure 11(a), $\bar{H}$ indicated that the target was well represented by its reference-group. However, $\bar{\rho}$ indicated a low correlation between the target and its reference-group. Hence, $\bar{\rho}$ and $\bar{H}$ were in conflict with each other. This made it a borderline case where it was a bit difficult to ascertain if the target was an outlier at the time of the creation of its reference-group. Interestingly, in Figure 11(b), while $\bar{\rho}$ indicated a strong relation between the target and its reference-group, $\bar{H}$ indicated otherwise. Both $\bar{\rho}$ and $\bar{H}$ were still in conflict with each other. However, the level of the return temperature of the target had fallen considerably compared to its reference-group. In this sense, the target was

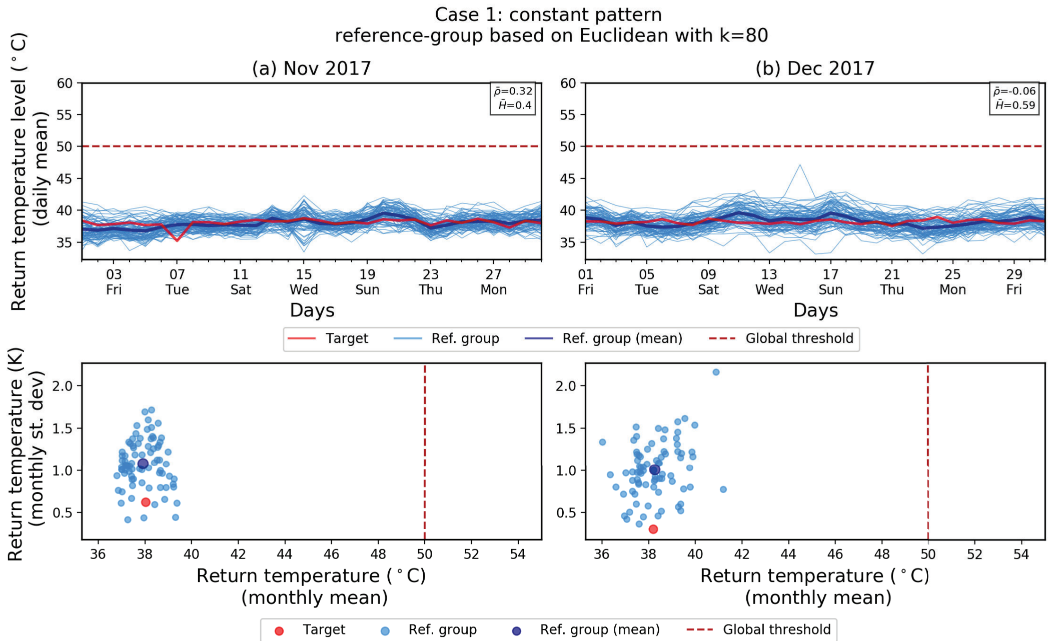

Figure 6: Example case of the "constant" pattern: The return temperature of the target stays at near constant levels compared to its reference-group.

no longer well represented by its reference-group. According to the DH expert, such decrease in return temperature could be due to maintenance service or changes in the configuration at the target substation or certain improvements in its associated building.

6. Discussion

In this section, we first discuss the reasons behind the selection of the return temperature variable used in this study. Next, we provide a reason behind the choice of using the data from the substations associated with the multi-dwelling building category. This is followed by a discussion on the basis of construction of the reference-groups and observations made therein. The results of comparison between global models, i.e., MCD and threshold, and the reference-group based approach are discussed next. Finally, we discuss the advantages and limitations of the reference-group approach.

The choice of return temperature for the analysis was based on the fact that it affects the overall operational efficiency of the entire DH network. Moreover, the return temperature is a good indicator of many problems in a substation. Furthermore, according to (Månsson et al., 2019), the analysis of return temperature levels is a widely used control check on the operational efficiency of substations in Sweden. The choice of multi-dwellings was based on the basis that it constitutes a significant proportion of more than 50% of total heat deliveries of the DH network investigated in this study.

Hence, the operational behavior of the return temperature from 778 substations associated with multi-dwellings in Helsingborg, Sweden, was studied using the reference-group based approach. The similarities among substations were measured on the basis of pointwise and distributional distances between their return temperature data. Analysis based on the stability proportion criterion was performed as shown in Figure 2 to determine which similarity measure and what k value is the best for constructing the reference-group. The results favored Euclidean distance with $k = 80$ to be the best available choice. Additionally, to study the effects of different similarity measures with various values of k on the detectability of target outliers, further analysis was conducted as shown in Table 1. Herein, it was observed that irrespective of the similarity measure used, there exists a trade-off in the detection of global and local outlier targets depending on the size k of the reference-group. A smaller k resulted in more local outlier targets. However, as k decreases, the possibility of false alarms increases.

The application of a global outlier detection model based on MCD with a contamination rate of 5% resulted in

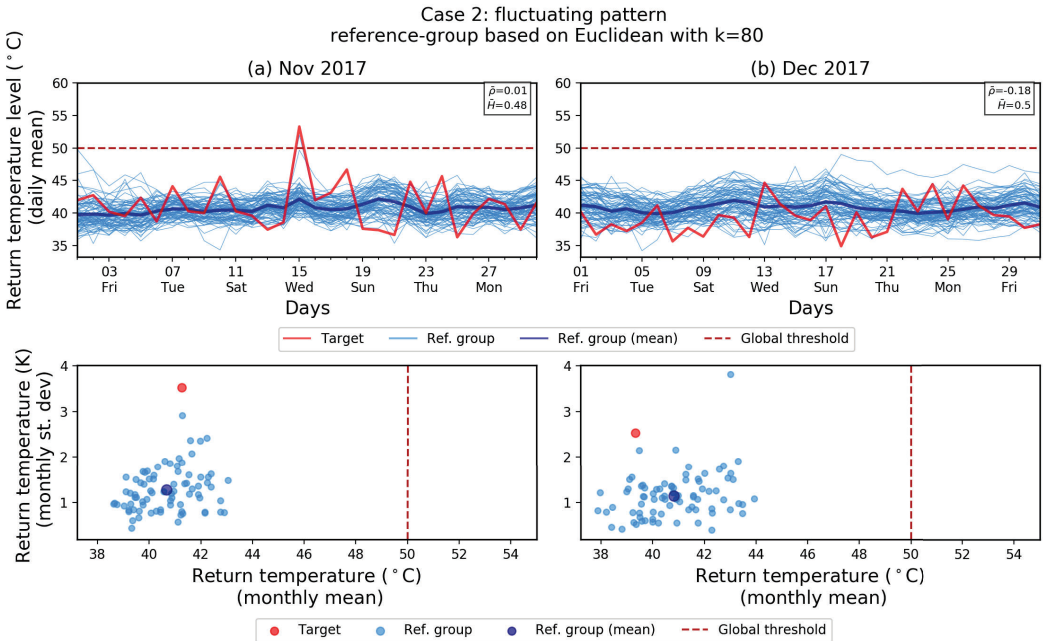

Figure 7: Example case of the "fluctuating" pattern: A possible malfunctioning target substation, which shows erratic behavior with high variability compared to its reference-group.

detection of a total of 39 global outlier targets among a total of 778 substations. Using a threshold of 50°C, only 6 global outlier targets were detected. In comparison, the reference-group approach under Euclidean distance with $k = 80$ detected a total of 77 outlier targets, 33 global and 44 local. This yielded an overall contamination rate of 10%. A total of 6 global outlier targets detected by the MCD model were missed by the reference-group approach. However, the 44 additional local outlier targets provided a possibility of detecting potentially problematic cases at a local level, albeit, at the cost of few false alarms. In this respect, a mere 5% increase in the outlier targets was justified. These local outlier targets could not have been detected by global models of outlier detection based on MCD and the threshold.

The main advantage of a reference-group based approach is that the reference operational behavior of any target substation in the network does not need to be predefined. Instead, the definition of normal, atypical or faulty operational behavior in a target substation is described relative to how its reference-group behaves. Thus, in effect, a reference-group acts as a local model of the target substation's operational behavior. It is therefore more efficient in detecting a deviating behavior compared to global models applied at the network level. Moreover, by relying on the operational behavior of the reference-group, this approach led to the identification of six most frequent patterns of deviating behavior in the return temperature of the substations. This can be useful to the DH utilities in deciding on where to look for to detect atypical and faulty operational behavior of a substation in a network.

The first limitation of the reference-group based approach is that not all target substations can be represented by reference-groups. For instance, it can be observed in Figure 3(b), that 68 targets among a total of 778 did not fulfill the adequacy criterion according to which the average Hellinger distance between the target and each member of its reference-group should be less than 0.4. Hence, these targets are not adequately represented by reference-groups at the episode of their construction, i.e., Nov'17. Such targets are either atypical or faulty to start with. In the former case, an individual model for the particular target may be required.

The second limitation is imposed by the DH network infrastructure itself. For the DH case, the reference-group based approach cannot be directly used by the individual substations to run any sort of automatic control mechanism when a problem is detected. The reason is that the focus of DH business has so far been on meeting the heat demand according to the customer's need. This is being achieved by controlling the supply temperature and pressure level via a centralized supervisory control and data actuation (SCADA) system, to deliver enough flow to the network in order to fulfill that requirement. However, most substations in operation currently are not configured for a SCADA system with regards to measurement, diagnosis, and load control (Gummerus, 2016). Over the last few years, improvements

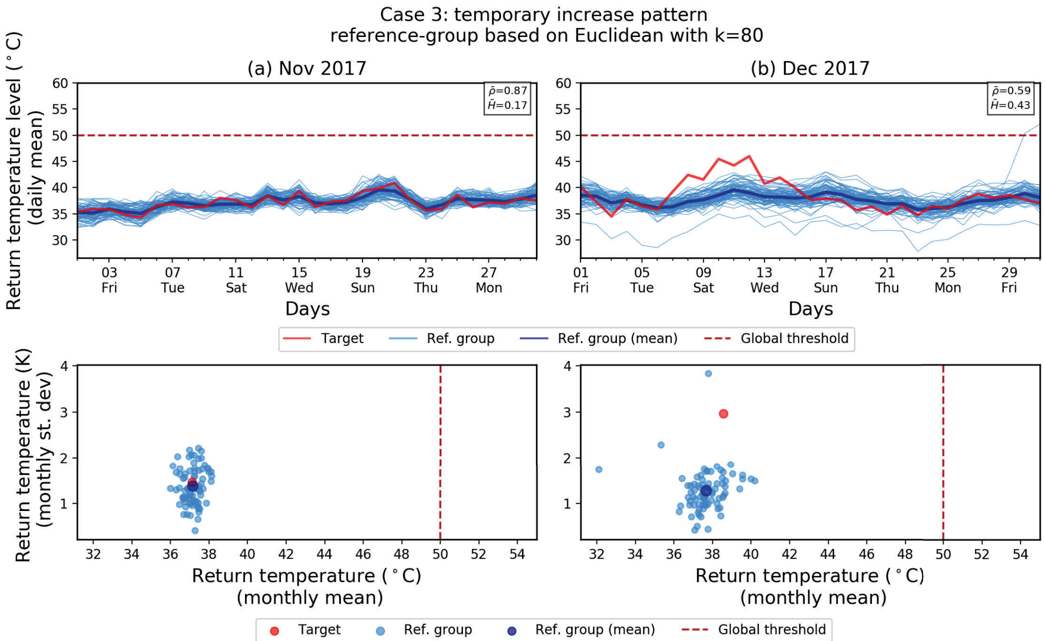

Figure 8: Example case of the "temporary increase" pattern: The return temperature of the target shows a considerable increase for about a week compared to its reference-group. A possible reason could be a control malfunction for short duration in Dec'17.

in technology and changes in energy regulations with regards to data collection have happened in the DH industry. This has enabled the possibility of improving the operational monitoring platform of DH networks, especially with regards to fault detection in substations. However, operational monitoring for control and optimization at the network level, which utilizes data and information on all the substations in a network, still remains a challenge (Gustafsson and Sandin, 2016). Thus, at present, the operational monitoring of substations, either via the central SCADA or condition monitoring (CM) systems, does not perform the analysis as suggested in this study. Therefore, the work done in this study can be programmed as a module for a DH utility for detecting atypical or faulty operational behavior of substations in its network.

7. Conclusion

Decrease in distribution temperatures is important for achieving the operational efficiency of a DH network. It is also a key requirement for a transition to 4GDH technology. Hence, understanding the operational behavior of distribution temperatures, especially the return temperature, is required at the substation level. Such analysis is now enabled due to digital or so-called "smart" meters installed at all the substations in Sweden. However, due to the constraints mentioned in Section 1, individual model at the substation level can be difficult to construct, while global models at the network level can be inefficient in terms of detecting deviating substations. Moreover, the common state of practice of using thresholds also has a limitation. That is, choosing a reference level for a threshold that is a compromise between a true alarm and a false alarm is usually unattainable. These constraints and limitations were addressed by formulating a reference-group based approach, which is described in Algorithm 1, see Section 3.5. There are three main advantages of using this approach. The first is that the reference operational behavior of any substation in a network does not need to be predefined. The second is that it provides a basis of what a substation's operational behavior should have been and what it is. In this respect, each system in the reference-group provides an evidence about a particular behavior during a particular time period. This can be very useful when a description of the normal, atypical or faulty operational behavior is unavailable. The third is that it leads to the detection of outlier substations that can be missed through the use of global models by providing for a more local context. These advantages have been demonstrated through the operational monitoring of the return temperature of 778 substations belonging to multi-

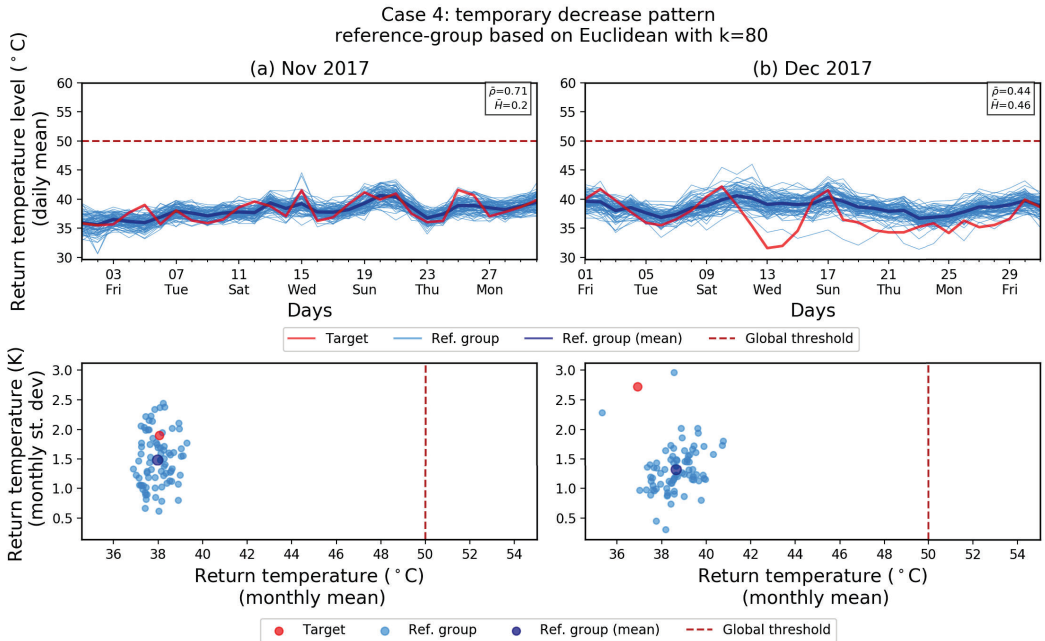

Figure 9: Example case of the "temporary decrease" pattern: A decrease in the return temperature of the target was observed for a few weeks of Dec'17 in comparison to its reference-group.

dwelling buildings located in Helsingborg, Sweden.

From a data-mining point of view, a reference-group based approach is useful not only in terms of isolating atypical and faulty operational behavior, but also in terms of interaction with the DH domain expert to determine why a particular operational behavior is occurring. Moreover, in the absence of a comprehensive fault-symptom datasets (Gunay et al., 2019), a reference-group based approach can provide a framework to label data of faulty substations. This can help in creating a knowledge base of faults in the DH industry in cases where they are not available.

While we have demonstrated this approach on the monitoring of the return temperature od substations, other relevant variables such as heating and flow rates together with the supply temperature can also be included. In fact, combining all the available relevant variables associated with substations is the best approach towards their operational monitoring. We will address this in the future. Moreover, the reference-group based approach may be applicable to other application domains where large-scale operational monitoring is required, such as in electricity utilities, solar and wind energy farms, factories with large fleet of manufacturing equipment or industrial robots, devices linked with Internet of Things (IoT).

We believe that in future smart energy systems, a system will not only require information on itself, but also knowledge about other comparable and related systems within the network. A reference-group based approach has the potential of enabling such information exchange.

Acknowledgments

This research is supported by the Swedish Knowledge Foundation (KK-stiftelsen) Grant No. 20160103. We would like to thank Öresundskraft, Sweden, for sharing their smart-meter dataset that was used in this study.

References

Bacher, P., Madsen, H., 2011. Identifying suitable models for the heat dynamics of buildings. Energy and Buildings 43, 1511–1522. URL: `http://www.sciencedirect.com/science/article/pii/S0378778811000491`, doi:`https://doi.org/10.1016/j.enbuild.2011.02.005`.

Bolton, R.J., Hand, D.J., 2001. Peer Group Analysis — Local Anomaly Detection in Longitudinal Data. Technical Report. Department of Mathematics, Imperial College London, UK.

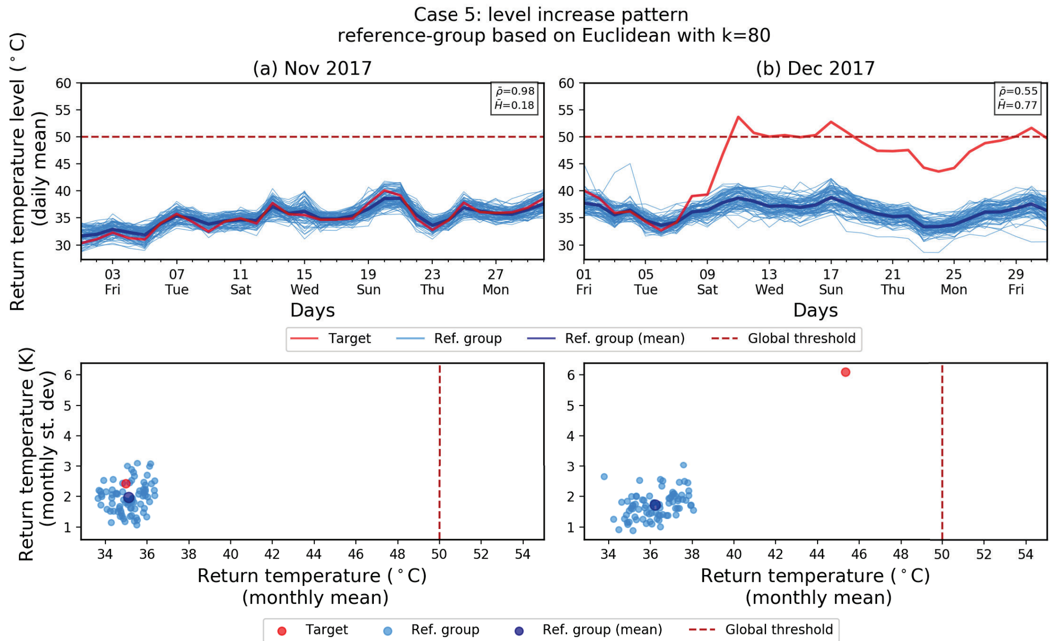

Figure 10: Example case of the "level increase" pattern: An abrupt increase in the return temperature of the target was observed in comparison to the reference-group in Dec'17. This could possibly be due to a fault in the flow control valve.

Breunig, M.M., Kriegel, H.P., Ng, R.T., Sander, J., 2000. LOF: Identifying density-based local outliers. SIGMOD Rec. 29, 93–104. URL: http://doi.acm.org/10.1145/335191.335388, doi:10.1145/335191.335388.

Byttner, S., Rögnvaldsson, T., Svensson, M., 2011. Consensus self-organized models for fault detection (COSMO). Engineering Applications of Artificial Intelligence 24, 833 – 839. URL: http://www.sciencedirect.com/science/article/pii/S0952197611000467, doi:https://doi.org/10.1016/j.engappai.2011.03.002.

Cai, B., Liu, Y., Fan, Q., Zhang, Y., Liu, Z., Yu, S., Ji, R., 2014. Multi-source information fusion based fault diagnosis of ground-source heat pump using bayesian network. Applied Energy 114, 1–9. URL: http://www.sciencedirect.com/science/article/pii/S0306261913007903, doi:https://doi.org/10.1016/j.apenergy.2013.09.043.

Cai, B., Shao, X., Liu, Y., Kong, X., Wang, H., Xu, H., Ge, W., 2019. Remaining useful life estimation of structure systems under the influence of multiple causes: Subsea pipelines as a case study. IEEE Transactions on Industrial Electronics doi:10.1109/TIE.2019.2931491.

Cover, T., 1982. This week's citation classic, current contents. 13.

Das, S., Matthews, B.L., Srivastava, A.N., Oza, N.C., 2010. Multiple kernel learning for heterogeneous anomaly detection: Algorithm and aviation safety case study, in: Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, New York, NY, USA. pp. 47–56. URL: http://doi.acm.org/10.1145/1835804.1835813, doi:10.1145/1835804.1835813.

Domingues, R., Filippone, M., Michiardi, P., Zouaoui, J., 2018. A comparative evaluation of outlier detection algorithms: Experiments and analyses. Pattern Recognition 74, 406–421. URL: http://www.sciencedirect.com/science/article/pii/S0031320317303916, doi:https://doi.org/10.1016/j.patcog.2017.09.037.

Emmott, A., Das, S., Dietterich, T., Fern, A., Wong, W.K., 2015. A meta-analysis of the anomaly detection problem. arXiv:1503.01158. [Accessed: 2019-12-05].

European Parliament, Council of the European Union, 2006. EU directive 2006/32/EC on energy end-use efficiency and energy services. URL: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32006L0032. [Accessed: 2019-08-22].

Fan, Y., Nowaczyk, S., Rögnvaldsson, T., 2015. Evaluation of self-organized approach for predicting compressor faults in a city bus fleet. Procedia Computer Science 53, 447 – 456. URL: http://www.sciencedirect.com/science/article/pii/S1877050915018256, doi:https://doi.org/10.1016/j.procs.2015.07.322. iNNS Conference on Big Data 2015 Program San Francisco, CA, USA 8-10 August 2015.

Fang, T., Lahdelma, R., 2016. Evaluation of a multiple linear regression model and sarima model in forecasting heat demand for district heating system. Applied Energy 179, 544–552. URL: http://www.sciencedirect.com/science/article/pii/S0306261916309217, doi:https://doi.org/10.1016/j.apenergy.2016.06.133.

Fontugne, R., Ortiz, J., Tremblay, N., Borgnat, P., Flandrin, P., Fukuda, K., Culler, D., Esaki, H., 2013. Strip, bind, and search: A method for identifying abnormal energy consumption in buildings, in: Proceedings of the 12th International Conference on Information Processing in Sensor Networks, ACM, New York, NY, USA. pp. 129–140. URL: http://doi.acm.org/10.1145/2461381.2461399, doi:10.1145/2461381.2461399.

Frederiksen, S., Werner, S., 2013. District heating and cooling, Studentlitteratur, Lund.

Gadd, H., Werner, S., 2013. Heat load patterns in district heating substations. Applied Energy 108, 176–183. URL: http://www.sciencedirect.

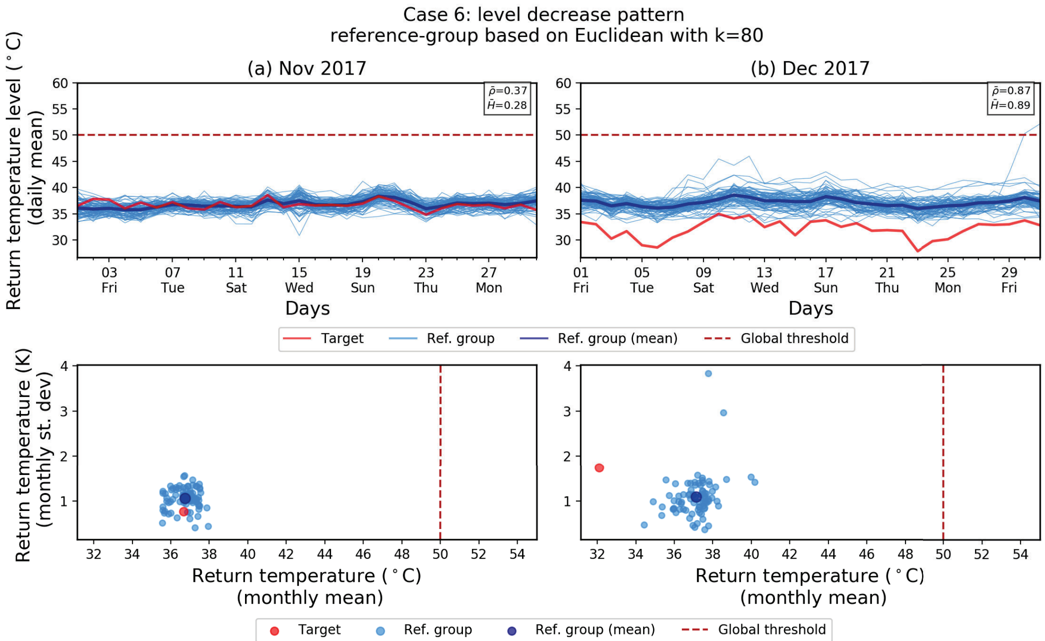

Figure 11: Example case of the "level decrease" pattern: A considerable decrease in observed in the return temperature of the target compared to its reference-group. This indicated towards possible improvements at the target substation or the building.

com/science/article/pii/S0306261913001803, doi:https://doi.org/10.1016/j.apenergy.2013.02.062.

Gadd, H., Werner, S., 2014. Achieving low return temperatures from district heating substations. Applied Energy 136, 59–67. URL: http://www.sciencedirect.com/science/article/pii/S0306261914009696, doi:https://doi.org/10.1016/j.apenergy.2014.09.022.

Gadd, H., Werner, S., 2015. Fault detection in district heating substations. Applied Energy 157, 51 – 59. URL: http://www.sciencedirect.com/science/article/pii/S0306261915009010, doi:https://doi.org/10.1016/j.apenergy.2015.07.061.

Gianniou, P., Liu, X., Heller, A., Nielsen, P.S., Rode, C., 2018. Clustering-based analysis for residential district heating data. Energy Conversion and Management 165, 840–850. URL: http://www.sciencedirect.com/science/article/pii/S019689041830236X, doi:https://doi.org/10.1016/j.enconman.2018.03.015.

Goldstein, M., Uchida, S., 2016. A comparative evaluation of unsupervised anomaly detection algorithms for multivariate data. PLOS ONE 11, 1–31. URL: https://doi.org/10.1371/journal.pone.0152173, doi:10.1371/journal.pone.0152173.

Grosswindhager, S., Voigt, A., Kozek, M., 2011. Online short-term forecast of system heat load in district heating networks, in: Proceedings of the 31st International Symposium on forecasting, Prague, Czech Republic.

Gummérus, P., 1989. Analysis of conventional substations in district heating systems. Ph.D. thesis. Chalmers University of Technology. Gothenburg, Sweden.

Gummerus, P., 2016. 10 - New developments in substations for district heating, in: Advanced District Heating and Cooling (DHC) Systems. Woodhead Publishing, Oxford. Woodhead Publishing Series in Energy, pp. 215–221. URL: http://www.sciencedirect.com/science/article/pii/B9781782423744000100, doi:https://doi.org/10.1016/B978-1-78242-374-4.00010-0.

Gunay, H.B., Shen, W., Newsham, G., 2019. Data analytics to improve building performance: A critical review. Automation in Construction 97, 96–109. URL: http://www.sciencedirect.com/science/article/pii/S0926580518305958, doi:https://doi.org/10.1016/j.autcon.2018.10.020.

Gustafsson, J., Sandin, F., 2016. 12 - District heating monitoring and control systems, in: Advanced District Heating and Cooling (DHC) Systems. Woodhead Publishing, Oxford. Woodhead Publishing Series in Energy, pp. 241–258. URL: http://www.sciencedirect.com/science/article/pii/B9781782423744000124, doi:https://doi.org/10.1016/B978-1-78242-374-4.00012-4.

Iyengar, S., Lee, S., Irwin, D., Shenoy, P., Weil, B., 2018. Watthome: A data-driven approach for energy efficiency analytics at city-scale, in: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, ACM, New York, NY, USA. pp. 396–405. URL: http://doi.acm.org/10.1145/3219819.3219825, doi:10.1145/3219819.3219825.

J. Pakanen, J. Hyvärinen, J.K., Ahonen, M., 1996. Fault diagnosis methods for district heating substations, Research Notes 1780. Technical Report. Technical Research Centre of Finland (VTT). Espoo.

Kim, W., Katipamula, S., 2018. A review of fault detection and diagnostics methods for building systems. Science and Technology for the Built Environment 24, 3–21. URL: https://doi.org/10.1080/23744731.2017.1318008, doi:10.1080/23744731.2017.1318008.

Lapira, E., 2012. Fault detection in a network of similar machines using clustering approach. Ph.D. thesis. University of Cincinnati.

Liu, F.T., Ting, K.M., Zhou, Z.H., 2012. Isolation-based anomaly detection. ACM Trans. Knowl. Discov. Data 6, 3:1–3:39. URL: http://doi.

acm.org/10.1145/2133360.2133363, doi:10.1145/2133360.2133363.

Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method. Methods 25, 402–408. URL: http://www.sciencedirect.com/science/article/pii/S1046202301912629, doi:https://doi.org/10.1006/meth.2001.1262.

Lund, H., Werner, S., Wiltshire, R., Svendsen, S., Thorsen, J.E., 2014. 4th generation district heating (4GDH): Integrating smart thermal grids into future sustainable energy systems. Energy 68, 1–11. URL: http://www.sciencedirect.com/science/article/pii/S0360544214002369, doi:https://doi.org/10.1016/j.energy.2014.02.089.

Ma, Z., Yan, R., Nord, N., 2017. A variation focused cluster analysis strategy to identify typical daily heating load profiles of higher education buildings. Energy 134, 90–102. URL: http://www.sciencedirect.com/science/article/pii/S0360544217309878, doi:https://doi.org/10.1016/j.energy.2017.05.191.

Månsson, S., Kallioniemi, P.O.J., Thern, M., Oevelen, T.V., Sernhed, K., 2019. Faults in district heating customer installations and ways to approach them: Experiences from Swedish utilities. Energy 180, 163–174. URL: http://www.sciencedirect.com/science/article/pii/S0360544219308606, doi:https://doi.org/10.1016/j.energy.2019.04.220.

Narayanaswamy, B., Balaji, B., Gupta, R., Agarwal, Y., 2014. Data driven investigation of faults in HVAC systems with model, cluster and compare (MCC), in: Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings, ACM, New York, NY, USA. pp. 50–59. URL: http://doi.acm.org/10.1145/2674061.2674067, doi:10.1145/2674061.2674067.

Nord, N., Nielsen, E.K.L., Kauko, H., Tereshchenko, T., 2018. Challenges and potentials for low-temperature district heating implementation in norway. Energy 151, 889 – 902. URL: http://www.sciencedirect.com/science/article/pii/S0360544218305036, doi:https://doi.org/10.1016/j.energy.2018.03.094.

Oza, N.C., Das, S., 2012. Anomaly detection in a fleet of systems. URL: https://c3.nasa.gov/dashlink/resources/607/. [Accessed: 2019-12-08].

Protić, M., Shamshirband, S., Petković, D., Abbasi, A., Kiah, M.L.M., Unar, J.A., Živković, L., Raos, M., 2015. Forecasting of consumers heat load in district heating systems using the support vector machine with a discrete wavelet transform algorithm. Energy 87, 343–351. URL: http://www.sciencedirect.com/science/article/pii/S0360544215005976, doi:https://doi.org/10.1016/j.energy.2015.04.109.

Räsänen, T., Ruuskanen, J., Kolehmainen, M., 2008. Reducing energy consumption by using self-organizing maps to create more personalized electricity use information. Applied Energy 85, 830 – 840.

Riksdagen, 2008. SFS 2008:263. Fjärrvärmelag. District heating act, Stockholm, Sweden.

Rousseeuw, P., Van Driessen, K., 1999. A fast algorithm for the minimum covariance determinant estimator. Technometrics 41, 212–223.

Sandin, F., Gustafsson, J., Delsing, J., 2013. Fault detection with hourly district energy data - probabilistic methods and heuristics for automated detection and ranking of anomalies. Technical Report. Svensk Fjärrvärme AB.

Schölkopf, B., Williamson, R., Smola, A., Shawe-Taylor, J., Platt, J., 1999. Support vector method for novelty detection, in: Proceedings of the 12th International Conference on Neural Information Processing Systems, MIT Press, Cambridge, MA, USA. pp. 582–588. URL: http://dl.acm.org/citation.cfm?id=3009657.3009740.

Sköldberg, H., Rydén, B., 2014. The heating market in Sweden — an overall picture. Technical Report. Värmemarknad, Sweden. URL: http://www.varmemarknad.se/pdf/The_heating_market_in_Sweden_141030.pdf.

Tureczek, A.M., Nielsen, P.S., Madsen, H., Brun, A., 2019. Clustering district heat exchange stations using smart meter consumption data. Energy and Buildings 182, 144 – 158. URL: http://www.sciencedirect.com/science/article/pii/S0378778818314725, doi:https://doi.org/10.1016/j.enbuild.2018.10.009.

Weston, D.J., Adams, N.M., Kim, Y., Hand, D.J., 2012. Fault mining using peer group analysis, in: Challenges at the Interface of Data Analysis, Computer Science, and Optimization, Springer Berlin Heidelberg, Berlin, Heidelberg. pp. 453–461.

Xue, P., Zhou, Z., Fang, X., Chen, X., Liu, L., Liu, Y., Liu, J., 2017. Fault detection and operation optimization in district heating substations based on data mining techniques. Applied Energy 205, 926–940. URL: http://www.sciencedirect.com/science/article/pii/S0306261917310401, doi:https://doi.org/10.1016/j.apenergy.2017.08.035.

Yao, R., Li, B., Liu, J., 2009. A theoretical adaptive model of thermal comfort – adaptive predicted mean vote (aPMV). Building and Environment 44, 2089–2096. URL: http://www.sciencedirect.com/science/article/pii/S0360132309000559, doi:https://doi.org/10.1016/j.buildenv.2009.02.014.