Energy and AI 17 (2024) 100393

Contents lists available at ScienceDirect

Energy and AI

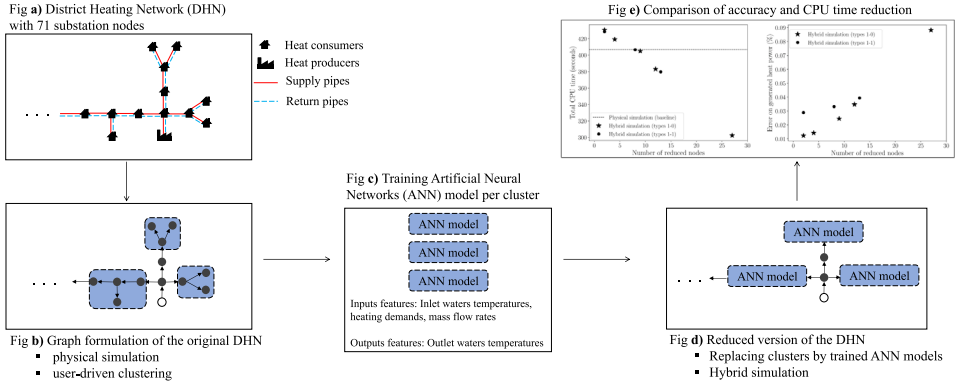
journal homepage: www.elsevier.com/locate/egyai

Topology reduction through machine learning to accelerate dynamic simulation of district heating

Dubon Rodrigue [a,b,*], Mohamed Tahar Mabrouk [a], Bastien Pasdeloup [b], Patrick Meyer [b], Bruno Lacarrière [a]

[a] IMT Atlantique, GEPEA, UMR CNRS 6144, 4 rue Alfred Kastler, Nantes, F-44307, France
[b] IMT Atlantique, Lab-STICC, UMR CNRS 6285, 655 Av. du Technopole, Brest, F-29238, France

GRAPHICAL ABSTRACT



HIGHLIGHTS

- Reducing dynamic simulations of District Heating Networks computational costs.
- Replacing predefined substation clusters with surrogate Artificial Neural Networks.
- Validation with wide diversity of clusters in terms of topology and heating demands.
- New hybrid simulation framework of the District Heating Networks.
- Reduction of simulation time while satisfyingly preserving produced heat powers.

ARTICLE INFO

Dataset link: https://github.com/drod-96/smart_clusters_v1.git, https://data.mendeley.com/datasets/77stj44drm/1

Keywords:
District heating network
Topology reduction
Artificial neural networks
Hybrid modeling

ABSTRACT

District heating networks (DHNs) provide an efficient heat distribution solution in urban areas, accomplished through interconnected and insulated pipes linking local heat sources to local consumers. This efficiency is further enhanced by the capacity of these networks to integrate renewable heat sources and thermal storage systems. However, integration of these systems adds complexity to the physical dynamics of the network, necessitating complex dynamic simulation models. These dynamic physical simulations are computationally expensive, limiting their adoption, particularly in large-scale networks. To address this challenge, we propose a methodology utilizing Artificial Neural Networks (ANNs) to reduce the computational time associated with the DHNs dynamic simulations. Our approach consists in replacing predefined clusters of substations

* Corresponding author at: IMT Atlantique, GEPEA, UMR CNRS 6144, 4 rue Alfred Kastler, Nantes, F-44307, France.
E-mail addresses: dubon.rodrigue@imt-atlantique.fr (D. Rodrigue), mohamed-tahar.mabrouk@imt-atlantique.fr (M.T. Mabrouk), bastien.pasdeloup@imt-atlantique.fr (B. Pasdeloup), patrick.meyer@imt-atlantique.fr (P. Meyer), bruno.lacarriere@imt-atlantique.fr (B. Lacarrière).

https://doi.org/10.1016/j.egyai.2024.100393
Received 16 April 2024; Received in revised form 8 June 2024; Accepted 26 June 2024
Available online 10 July 2024

Graph-based formulation

within the DHNs with trained surrogate ANNs models, effectively transforming these clusters into single nodes. This creates a hybrid simulation framework combining the predictions of the ANNs models with the accurate physical simulations of remaining substation nodes and pipes. We evaluate different architectures of Artificial Neural Network on diverse clusters from four synthetic DHNs with realistic heating demands. Results demonstrate that ANNs effectively learn cluster dynamics irrespective of topology or heating demand levels. Through our experiments, we achieved a 27% reduction in simulation time by replacing 39% of consumer nodes while maintaining acceptable accuracy in preserving the generated heat powers by sources.

1. Introduction

District Heating Networks (DHNs) are urban-scale heat distribution systems. They offer an efficient way to transport heat energy from local heat sources to local heat consumers through a network of insulated pipes. This efficiency is further enhanced by the integration of decentralized heat sources, including renewable sources [1], and thermal storage systems. While improving the efficiency, the integration of the systems also increases the complexity of the physical dynamics of the DHNs. This complexity necessitates the use of dynamic based simulation models. However, these models are computational expensive, especially for large-scale DHNs. Two main factors contribute to such high computational cost. First, these simulation models require a complete topological representation of the entire DHN. Second, the spatial discretization of the pipes leads to computationally demanding matrix inversions. Developing methods to reduce these computational costs is therefore crucial to enable wider adoption of DHNs dynamic simulation models.

To address this computational cost, three main categories of methods have emerged. The first category focuses on the use of simplified physical modeling [2] or pseudo-dynamic approach [3]. However, it comes at the expense of accuracy in capturing the complex dynamic behavior of the network. The second category includes Model Order Reduction (MOR) approaches which tackle topological complexity by projecting the physical state description onto a lower-dimensional space. Techniques such as proper orthogonal decomposition [4] and Galerkin projection [5] have been explored in this context. Rein et al. [6] noted that achieving substantial gains often requires a significant reduction in the dimensions of the space, sacrificing physical details. A recent approach in this category involving network partitioning starts by dividing the network into clusters (subsets of the substation nodes) and applying MOR at the cluster level [7]. The third and last category includes topology reduction based methods which simplify the network topology itself. This approach also divides the network into clusters but, in contrast to MOR approach, focuses on simplifying the topology within each cluster, without altering the underlying physical equations [8].

Topology or grid reduction methods aim to simplify the topology of the clusters by removing specific connections (pipes) or nodes (substations). In the context of aggregation frameworks, the entire cluster is replaced by a single node. For DHN applications, this translates into substituting the original cluster, which includes its substations and connecting pipes, with a single equivalent node. The key challenge lies in preserving the physical equivalence between the original and reduced networks. This necessitates maintaining identical thermal power distribution in the non-aggregated regions of both networks versions. It ensures also similar generated thermal power by the heat sources. Two major DHN aggregation methods are present in the literature: the German [9] and Danish [10] methods. To conserve the thermal equivalency, these methods empirically adjust the properties of the pipes of the reduced network, allowing to have the same transport delays and mass flow rates as the original network. These methods have been successfully applied to operational optimization tasks [11], demonstrating promising results in conserving the thermal power generated from heat sources and reducing the simulation time. However, limitations have been identified, including the requirement for steady-state conditions and the restriction to tree-shaped clusters [11]. A recent work by Kane et al. [12] attempts to address the tree-shaped limitation by incorporating clusters with loop connectivity. However, this approach still relies on steady-state physical models. Besides, it is based on assumption that the mass flow rate of the water through the pipes evolves proportionally, which may not reflect real-world DHN operations. Although these aggregation methods show promising results, these limitations hinder their widespread adoption.

The past decades have witnessed the remarkable performance of Machine Learning (ML) models across various engineering disciplines. The energy sector is no exception, although its adoption is still in its early stages [13]. Unsupervised ML techniques, such as k-means [14] and k-shape [15], have been explored to partition DHN into clusters of substations that exhibit similar consumption patterns. Although these methods have facilitated the application of control strategies at the cluster level [15], they lack a mechanism to replace identified clusters. A potentially more effective approach involves employing supervised ML models to learn and replicate the physical dynamics of the clusters. These trained models could then serve as surrogate models of the clusters within the aggregation framework. In particular, there is a lack of research investigating such ML-based aggregation approaches specifically for DHNs. Artificial Neural Networks (ANNs) have indeed demonstrated exceptional capabilities in learning complex non-linear relationships [16], frequently observed in physical phenomena. Different architectures of ANNs models, including Multi-Layer Perceptrons (MLP) [17], Convolutional Neural Networks (CNN) [18], and Recurrent Neural Networks (RNN) [19], have been applied to thermal networks. However, prior studies have primarily focused on: replacing the entire network with a "black-box" model [20] acting as surrogate model of the simulation model, forecasting demand power [19], or estimating specific physical properties such as thermal conductivity [18]. Our research identified no existing methodology for learning and substituting substation clusters within the DHNs. This cluster-level substitution approach offers a more efficient alternative to learning and replacing the entire network. It requires less data and minimizes potential error sources.

This paper addresses the identified knowledge gaps by proposing an aggregation methodology to reduce the computational costs related to the dynamic simulations of the DHN. Our approach focuses on replacing predefined clusters of substations with trained ANNs models, acting as surrogate models, in an aggregation framework. This methodology uses dynamic physical simulations of the DHN. Also, no prior assumptions on the mass flow rate of the water and the topology of the clusters are required. Furthermore, a hybrid simulation framework is introduced which couples the surrogate ANNs models predictions with the physical simulation. The remainder of this paper is structured as follows. Section 2 presents the physical modeling of the DHN and details our aggregation process. In Section 3, we explore the different ANNs architectures considered for this work. To validate the effectiveness of the proposed methodology, Section 4 describes the experiments carried out, including details of the studied DHNs and the selected clusters. Section 5 compares the performance of the various ANNs architectures and evaluates the computational gains achieved through the reduced DHN hybrid simulation. Finally, Section 6 summarizes the key findings of the study and provides a comprehensive discussion of their implications.

2. Methodology

In line with the node method [21], the DHN is viewed as a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ and $\mathcal{E}$ represent the sets of nodes and edges within the DHN, respectively.

2.1. Notations

In our work, nodes are identified by letters (*e.g.*, v) or by an integer (*e.g.*, 1). While prior DHN modeling literature may use terms like "edge", "boundary", and "pipe" interchangeably to denote connections, this paper adopts a distinct definition. In our work, an "edge" represents a pair of parallel pipes with identical properties but opposite flow directions: a supply pipe and a return pipe. An edge denoted as (u, v) indicates a connection between nodes u and v, where the supply pipe flows from u to v and the return pipe flows in the opposite direction. Pipes oriented towards a node are termed incoming, while those oriented away are termed outgoing. Time-varying values are italic-bolded, while static values are italic-normal. Subscripts s and r denote the temperatures of the supply and return water, respectively. Additionally, the superscripts *in* and *out* indicate the temperatures of the water at the inlet and outlet of the pipes (supply and return), respectively. We denote the sets of consumer and source nodes as $\mathcal{V}_C$ and $\mathcal{V}_S$, respectively, with their disjoint union corresponding to the set of all the nodes within the DHN (*i.e.*, $\mathcal{V} = \mathcal{V}_C \sqcup \mathcal{V}_S$).

2.2. Dynamic physical modeling of the DHN

Here, we introduce the dynamic physical modeling used to simulate the DHN which leverages nonlinear partial derivative equations to effectively capture the behavior of the DHN. The direction of the edges arbitrarily follows the direction of the supply pipes. This simulation model is composed of a thermal model and a hydraulic model. The nodes in the DHN represent either substations (heat exchangers) or simply flow mixing locations, with the supply $\boldsymbol{T_s}$ and return $\boldsymbol{T_r}$ temperatures.

2.2.1. Thermal model

Substation nodes can act as sources (injecting heat) or consumers (consuming heat). Prosumers (combined source-consumer nodes) and storage nodes are not included in this work. Heat injection and consumption are governed, respectively, by the Eqs. (1) and (2) where c_p is the specific heat capacity of the water. Sources supply heat power $\boldsymbol{P}$ into the network by injecting hot supply water with supply mass flow rates $\dot{\boldsymbol{m}}_s$ at generation supply temperatures $\boldsymbol{T}_{ss}$ (see Fig. 1(a)). Consumers extract heat power to satisfy their heating demands $\boldsymbol{D}$, achieved by consuming supply water with consumption mass flow rates $\dot{\boldsymbol{m}}_c$ and returning it at a fixed primary return temperature T_{rc} (see Fig. 1(b)). Mixing nodes simply mix incoming flows. $\mathcal{E}_v^+$ (Eq. (3)) and $\mathcal{E}_v^+$ (Eq. (4)) contain respectively the subsets of edges with incoming and outgoing supply flows of a node v, respectively, which depend on the direction of the edges in $\mathcal{E}$ and the signs of the traversing mass flow rate $\dot{\boldsymbol{m}}$ of the water. Based on these subsets, we can derive the incoming and outgoing return flows of the node v. Eqs. (6) and (5) ensure enthalpy conservation at each node for both the supply and return sides. Heat transport through the pipes (supply and return) is modeled by the heat transport equation presented in Eq. (8) where $\boldsymbol{T}$ indicates the temperatures of the water through the pipes (supply and return). Pipes characteristics values are presented in Table 1. ρ represents the density of the water. The heat transport equation breaks down into three key components. First, the change in temperature over time, captured by the partial derivative, reflects the heat accumulation within the pipes. Second, the partial derivative with respect to the direction of flow (*i.e.*, x-axis) captures the transport of the heat by convection along the pipes. The values of the mass flow rates of the water $\dot{\boldsymbol{m}}_{(u,v)}$ are identical for supply and return pipes along each edge (u, v) due to mass conservation.

Table 1
Pipes characteristics and corresponding physical units.

Parameters	Physical name	Physical unit
d	Diameter	[m]
h	Thermal convective loss coefficient	[W.K^{-1}.m^{-2}]
l	Length	[m]
k	Pressure resistance coefficient	m.s^2.kg^{-2}

The upwind scheme and finite-volume discretization solve Eq. (8) with boundary conditions imposed by enthalpy conservation (Eqs. (5)–(6)) and numerical continuity (Eq. (7)). According to such a resolution approach, the required computation cost scales with the size of the network ($|\mathcal{E}|, |\mathcal{V}|$), the number of discretized finite volumes N_{volumes} and the number of time steps N_{steps}. It highlights the need for the DHN topology reduction.

$$\forall s \in \mathcal{V}_S,\ \boldsymbol{P}_s = \dot{\boldsymbol{m}}_{s_s} c_p (\boldsymbol{T}_{ss_s} - \boldsymbol{T}_{r_s}) \tag{1}$$

$$\forall c \in \mathcal{V}_C,\ \boldsymbol{D}_c = \dot{\boldsymbol{m}}_{c_c} c_p (\boldsymbol{T}_{s_c} - T_{rc_c}) \tag{2}$$

$$\forall v \in \mathcal{V},\ \mathcal{E}_v^+ = \left\{(u, v) \in \mathcal{E}, \dot{\boldsymbol{m}}_{(u,v)} > 0\right\} \cup \left\{(v, w) \in \mathcal{E}, \dot{\boldsymbol{m}}_{(v,w)} < 0\right\} \tag{3}$$

$$\forall v \in \mathcal{V},\ \mathcal{E}_v^- = \left\{(v, w) \in \mathcal{E}, \dot{\boldsymbol{m}}_{(v,w)} > 0\right\} \cup \left\{(u, v) \in \mathcal{E}, \dot{\boldsymbol{m}}_{(u,v)} < 0\right\} \tag{4}$$

$$\forall v \in \mathcal{V},\ \sum_{(u,v)\in\mathcal{E}_v^+} |\dot{\boldsymbol{m}}_{(u,v)}| \boldsymbol{T}^{\,out}_{s_{(u,v)}} + \dot{\boldsymbol{m}}_{s_v} \boldsymbol{T}_{ss_v} = \sum_{(v,w)\in\mathcal{E}_v^-} |\dot{\boldsymbol{m}}_{(v,w)}| \boldsymbol{T}_{s_v} \tag{5}$$

$$\forall v \in \mathcal{V},\ \sum_{(v,w)\in\mathcal{E}_v^-} |\dot{\boldsymbol{m}}_{(v,w)}| \boldsymbol{T}^{\,out}_{r_{(v,w)}} + \dot{\boldsymbol{m}}_{c_v} T_{rc_v} = \sum_{(u,v)\in\mathcal{E}_v^+} |\dot{\boldsymbol{m}}_{(u,v)}| \boldsymbol{T}_{r_v} \tag{6}$$

$$\forall (u, v) \in \mathcal{E},\ \boldsymbol{T}_{s_v} = \boldsymbol{T}^{\,in}_{s_{(u,v)}} \text{ and } \boldsymbol{T}_{r_v} = \boldsymbol{T}^{\,in}_{r_{(v,u)}} \tag{7}$$

$$\forall (u, v) \in \mathcal{E},\ \frac{\partial \boldsymbol{T}_{(u,v)}}{\partial t} = -\frac{4\dot{\boldsymbol{m}}_{(u,v)}}{\pi \rho d^2_{(u,v)}} \frac{\partial \boldsymbol{T}_{(u,v)}}{\partial x} - \frac{4h_{(u,v)}}{\rho c_p d_{(u,v)}} \left(\boldsymbol{T}_{(u,v)} - \boldsymbol{T}_{\text{ground}}\right) \tag{8}$$

2.2.2. Hydraulic model

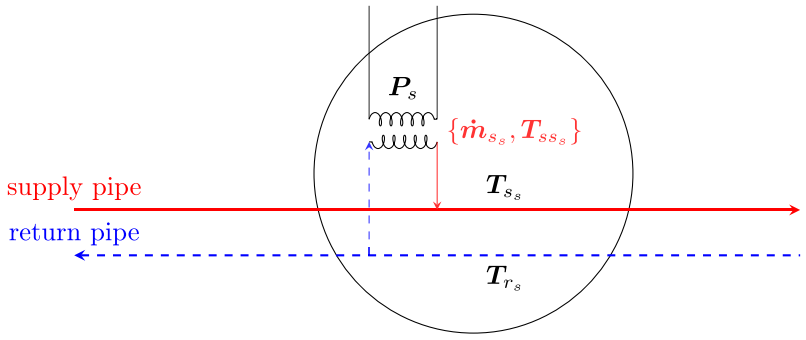
The hydraulic model incorporates two key aspects: nodal mass balance and pressure drops within the pipes. Eq. (9) enforces mass conservation at every node and time step t. Pressure drops through the pipes are modeled by pressure head $\boldsymbol{H}$ losses between adjacent nodes, governed by Eq. (10). This equation is derived from the Darcy–Weisbach pressure loss equation [22]. The pressure-resistance coefficient k (see Table 1) depends on both the mass flow rate of the water through the pipes and their surface roughness [23].

$$\forall t,\ \forall v \in \mathcal{V},\ \dot{\boldsymbol{m}}^t_{s_v} + \sum_{(u,v)\in\mathcal{E}_v^+} \dot{\boldsymbol{m}}^t_{(u,v)} = \dot{\boldsymbol{m}}^t_{c_v} + \sum_{(v,w)\in\mathcal{E}_v^-} \dot{\boldsymbol{m}}^t_{(v,w)} \tag{9}$$

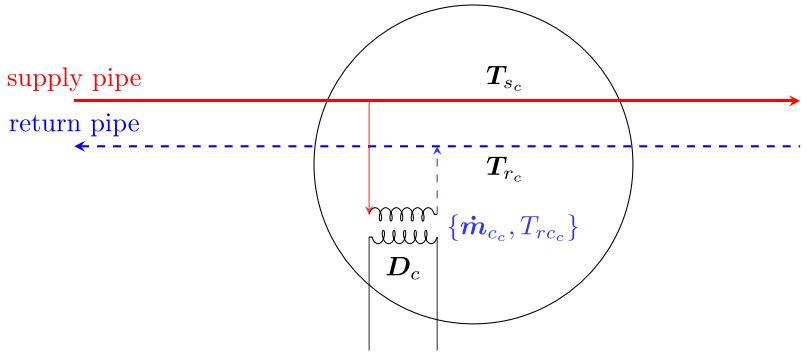
$$\forall (u, v) \in \mathcal{E},\ \boldsymbol{H}_u - \boldsymbol{H}_v = k_{(u,v)} \dot{\boldsymbol{m}}_{(u,v)} |\dot{\boldsymbol{m}}_{(u,v)}| \tag{10}$$

2.3. ML aided topology reduction

The physics-based modeling provides a baseline and data to train the proposed learning and aggregation framework. This framework aims to simplify the DHN topology by replacing predefined consumer clusters with surrogate ANNs models. These ANNs models are trained to mimic the underlying physical dynamics of the clusters and act as <<*twin nodes*>> of the original clusters. In this work, we only consider clusters of consumer nodes, excluding source nodes.



(a) A source node s with produced heat P_s



(b) A consumer node c with heating demand D_c

Fig. 1. Illustration of substation node modeling, for sources 1(b) and consumers 1(a).

2.3.1. Graph formulation of the DHN

We extend the node method graph-based representation of the DHN by assigning multivariate time series attributes to the nodes and the edges. Attributes of the nodes include generated heat power (zero for consumers), heating demands, supply and return temperatures forming a four-feature vector $\boldsymbol{X}_v = \left\{\boldsymbol{P}_v, \boldsymbol{D}_v, \boldsymbol{T}_{s_v}, \boldsymbol{T}_{r_v}\right\}$ for every node $v \in \mathcal{V}$. In this graph-based representation, edge orientations follow the supply flows. We assign to each edge $(u,v) \in \mathcal{E}$ a multivariate time series weight $\boldsymbol{W}_{(u,v)} = \left\{\boldsymbol{W}^{in}_{s_{(u,v)}}, \boldsymbol{W}^{in}_{r_{(u,v)}}, \boldsymbol{W}^{out}_{s_{(u,v)}}, \boldsymbol{W}^{out}_{r_{(u,v)}}\right\}$, which are respectively the inlet and outlet temperatures of the corresponding supply and return pipes (*i.e.*, based on the directions of the flows). These values are defined in Eq. (11). Inlet temperatures are multiplied by a time-dependent coefficient $\boldsymbol{\omega}_{(u,v)}$ accounting for the characteristics of the pipes and mass flow rates. This coefficient is derived from a steady-state thermal loss estimation method [24]. We use the same $\boldsymbol{\omega}_{(u,v)}$ for both supply and return due to the similarities of the pipes and mass conservation.

$$\forall (u,v) \in \mathcal{E}, \begin{cases} \boldsymbol{\omega}_{(u,v)} = \exp\left(-\frac{d_{(u,v)} l_{(u,v)} h_{(u,v)}}{c_p \dot{\boldsymbol{m}}_{(u,v)}}\right) \\ \boldsymbol{W}^{in}_{r_{(u,v)}} = \boldsymbol{\omega}_{(u,v)} \odot \boldsymbol{T}^{in}_{r_{(u,v)}} \\ \boldsymbol{W}^{in}_{s_{(u,v)}} = \boldsymbol{\omega}_{(u,v)} \odot \boldsymbol{T}^{in}_{s_{(u,v)}} \\ \boldsymbol{W}^{out}_{r_{(u,v)}} = \boldsymbol{T}^{out}_{r_{(u,v)}} \\ \boldsymbol{W}^{out}_{s_{(u,v)}} = \boldsymbol{T}^{out}_{s_{(u,v)}} \end{cases} \tag{11}$$

2.3.2. Learning and aggregation framework

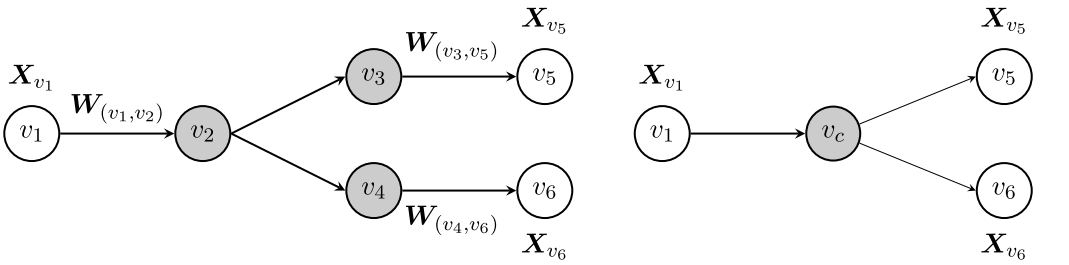
Given a cluster (C), its surrogate ANNs model predicts the temperatures of outgoing flows of the water at the outlets of the pipes (supply and return), collectively designed by the output features $\boldsymbol{Y}$. The ANNs model considers the temperatures of incoming flows of the water at the inlets of the pipes and the heating demands of the consumer nodes composing the replaced cluster, collectively designed as input features $\boldsymbol{Z}$. Crucially, identifying the input and output features of the surrogate models of the aggregated clusters requires prior knowledge of the external edges of these clusters. These are the edges between a cluster and adjacent nodes outside the cluster. Flows of the water through these external edges, entering or exiting the cluster. In other words, external edges represent the interface between a cluster and the rest of the network. To illustrate this aggregation framework, consider the example in Fig. 2. The cluster (v_1, v_2, v_3) of size three is aggregated into its twin node v_c. The physical states of the adjacent nodes (v_1, v_5, v_6) outside the cluster, connected through external edges $((v_1, v_2), (v_3, v_5), (v_4, v_6))$, are preserved. Thermal equivalency between the original network (Fig. 2(a)) and the reduced network (Fig. 2(b)) is ensured by learning and predicting outgoing flows temperatures $\boldsymbol{Y}$ form the inputs $\boldsymbol{Z}$, defined in Eqs. (12) and (13) respectively. Notably, the ANNs models do not use any internal temperatures of the original cluster as input features. Indeed, during simulations of the reduced network, these internal temperatures will not be readily available. The proposed learning and aggregation process is adaptable to any cluster configuration. By construction, the number of output features is directly related to the number of external edges in the original cluster and the directions of the flows of the water. Crucially, our framework assumes that no flow reversal occurs in the external edges through the simulation. Currently, we categorize clusters based on their type, defined by the combination of incoming and outgoing edges. This type directly reflects the number of outgoing flows and the temperatures that the surrogate ANNs model needs to predict. For example, the cluster in Fig. 2 is classified as type 1–2, indicating one incoming and two outgoing edges.

$$\boldsymbol{Y} = \left\{\boldsymbol{W}^{out}_{r_{(v_1,v_2)}}, \boldsymbol{W}^{out}_{s_{(v_3,v_5)}}, \boldsymbol{W}^{out}_{s_{(v_4,v_6)}}\right\} \tag{12}$$

$$\boldsymbol{Z} = \left\{\boldsymbol{D}_{v_2}, \boldsymbol{D}_{v_3}, \boldsymbol{D}_{v_4}, \boldsymbol{W}^{in}_{s_{(v_1,v_2)}}, \boldsymbol{W}^{in}_{r_{(v_3,v_5)}}, \boldsymbol{W}^{in}_{r_{(v_4,v_6)}}\right\} \tag{13}$$

2.4. Hybrid simulation

Replacing the clusters with single twin nodes reduces the topology of the DHN by reducing the number of physical pipes and nodes. To



(a) Original network with consumer nodes cluster $\{v_2, v_3, v_4\}$ in grey.

(b) Reduced network with cluster replaced by its twin node v_c.

Fig. 2. Illustration of the proposed aggregation process on an example cluster.

simulate such reduced networks, the ANNs models, associated to these twin nodes, communicate with the physical simulation. In a more detailed way, the physical simulation of the remaining physical pipes and nodes give the input features of the ANNs models whose predictions are fed back into the physical simulation to finish the network overall simulation. This combined physical and ANNs models predictions is referred to as hybrid simulation. This proposed hybrid framework is adaptable and can be applied to any simulation model. It necessitates and involves the following data pipelines: (a) collecting the required data at each time step from the simulation models, (b) preprocessing the collected data for use as inputs to the ANNs, (c) using the ANNs to make predictions, and (d) feeding the predicted values back into the DHN simulations to finalize the process. In our paper, we focus on leveraging data from DHN simulation models, although this approach can be adapted to use field measurements. To evaluate the performance of the hybrid framework, we concentrate on two aspects: accuracy and computational efficiency. We measure efficiency by the reduction in simulation CPU time. To assess accuracy, we define two physics-based metrics. First, we measure the ability of the hybrid simulation to conserve the thermal loads at the interfaces of the twin nodes, *i.e.*, replaced clusters. In this work, the thermal loads refer to available heat powers, upstream or downstream of the clusters. Our second accuracy metric consists in measuring the ability of the hybrid simulation to conserve the produced heat powers by the sources. Both metrics are measured in absolute relative errors (%), averaged over time.

3. ANNs models

To effectively learn the complex physical dynamics of the clusters, we evaluate three compact ANNs architectures: Gated recurrent neural networks (GRU-NN), convolutional neural networks (CNN), and multi-layer perceptrons (MLP). The following sections detail these architectures, all selected to balance precision with computational efficiency.

3.1. GRU–NN architectures

RNN based models are indeed suitable for time series data. Gated Recurrent Unit (GRU) [25] has been recently introduced to address the vanishing gradients, a challenge in training vanilla RNN models. These cells allow to capture long time dependency in temporal data. In this work, we explore many-to-one RNN architecture relying on GRU cells. To capture the transport delay time through the clusters, we use input sequences of length 60, *i.e.*, prediction horizon, containing input features between time steps $t-60$ to t, to predict the output features at time step t. Two versions of the GRU-NN architectures, shown in Fig. 3, are compared. GRU cells have 20 units and use the hyperbolic tangent as an activation function while dense layers with 60 units use Rectified Linear Unit activation function.

3.2. CNN architectures

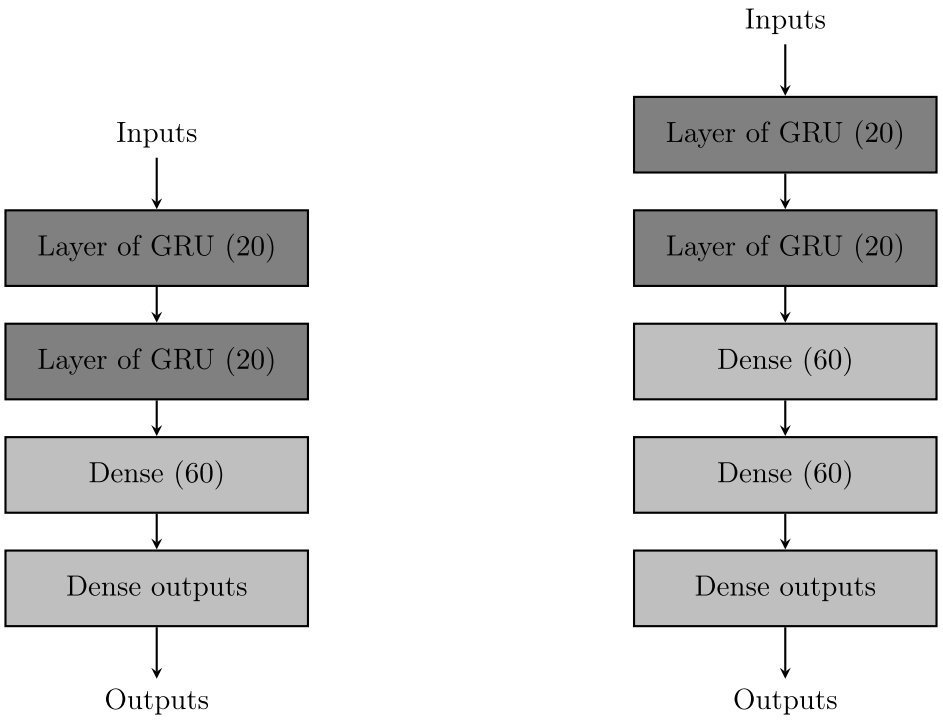
For efficiency, 1D CNNs, also known as temporal convolutions, are considered. 1D convolution, designed for 1D signals, requires less computation than the 2D or 3D counterparts [26]. Similar to GRU-NNs, CNN models use input sequences of length δ_t, viewed as matrices, to predict output features at time t. Also, we compare two versions of the CNN architectures, shown in Fig. 4. Valid padding, 1 stride, and scaled exponential linear unit (selu) activation [27] are considered by the convolution layers. Average pooling layers are placed after each convolution layer without strides.

3.3. MLP architecture

While requiring computational load less than RNNs and CNNs, MLP based models traditionally handle single time steps predictions. We explore two approaches for incorporating time series: time-to-time prediction and flattened sequences. Both are depicted in Fig. 5. Flattening the sequences, however, suffers from high dimensionality of the input features, leading to large size of models. We compare these input features engineering for the same MLP architecture.

3.4. Training procedures

Let $\boldsymbol{Z} \in \mathbb{R}^{\mathrm{F}\times N_{\text{steps}}}$ denote the input time-series and $\boldsymbol{Y} \in \mathbb{R}^{\mathrm{O}\times N_{\text{steps}}}$ the output time-series for a given cluster, where F and O represent the number of input and output features, respectively. Both $\boldsymbol{Z}_f \in \mathbb{R}^{N_{\text{steps}}}$ and $\boldsymbol{Y}_o \in \mathbb{R}^{N_{\text{steps}}}$ (feature vectors of the input and output time series) are Min–Max normalized, along the time axis. The normalized $\bar{\boldsymbol{Z}}$ is then partitioned into sequences $\left\{z_i \in \mathbb{R}^{\mathrm{F}\times\delta_t}\right\}_{i=1}^{n}$ of length δ_t along the time-axis. Similarly, $\bar{\boldsymbol{Y}}$ is partitioned, but only the last point of each sequence is kept, resulting in output sets $\left\{y_i \in \mathbb{R}^{\mathrm{O}\times 1}\right\}_{i=1}^{n}$. n indicates the total number of data sets and corresponds to $N_{\text{steps}} - \delta_t + 1$. We temporally split the data sets into training and test sets, representing 80% and 20% of the data sets respectively. Then the test and training sets are shuffled along the sets. During the training phase, 20% of the training sets are used as validation sets. The ANNs based models learn to minimize the mean absolute error (MAE) on the training sets (Eq. (14)) between the ground-truth values y_i from the physical simulation and the predicted values $\tilde{y}_i$. n_t is the number of sets in the training batch. Reported performances of our ANNs models in this paper are measured on the test sets (in °C). We train the models on a maximum number of epochs 100 using Adam optimizer and batch size of 100. An early-stop callback monitors the improvement of the training metrics (*i.e.*, MAE) in the validation sets and prematurely stops the training loop if no improvement has been observed during the last 20 epochs. The learning rate starts from 10^{-3} and sequentially decreases by a factor of 0.1 every 20 epochs to 10^{-5}. Indeed, we have noticed that these step decreases of the learning rate improve convergence of the training process. These training callbacks serve as safeguards against overfitting the ANNs models to the training data, ensuring they can generalize



(a) GRU-NN v1 architecture (b) GRU-NN v2 architecture

Fig. 3. The considered GRU-NN architectures versions. The number of units per GRU and Dense layer are shown in the cases.

to the wide range of physical data variations encountered over the simulation years.

$$\frac{1}{n_t}\sum_{i=1}^{n_t}|y_i - \tilde{y}_i| \tag{14}$$

4. Experiments

To measure the performance and generalizability of our proposed aggregation framework, we apply our learning methodology to a large diversity of clusters. These clusters are selected from four different synthetic DHNs, which serve as our case study DHNs in this work (see Fig. B.1). Each cluster is learned independently using the proposed ANNs architectures, regardless of the specific topology of the DHN it originated from. The DHNs with fixed topology are simulated using the previously described physical model with realistic heating demands, generated using a heating law (more details in Appendix A). The simulation data ranges from January to Mid-June of 2019 (5.5 months) with a 1-minute time step, resulting in $N_{\text{steps}} = 224,540$. The considered outdoor temperatures are obtained from the NASA power publicly available database [28]. The DHNs have two distributed heat sources capable of supplying all heating demands of the consumers at all time steps. The generation supply temperatures of the sources follow a linear law in function of the outdoor temperatures (see Appendix A). Additionally, $N_{\text{volumes}} = 200$ is considered for all pipes. The training and testing data for the ANNs models are derived from the simulations of our four case-study DHNs, employing the data pipeline outlined in Section 3.4.

4.1. User-driven clustering

To comprehensively evaluate the performance of our methodology across diverse cluster configurations, exhaustively considering all possible clusters in a DHN is impractical due to the exponential growth in the number of potential clusters within a graph. Therefore, we focus on capturing a significant range of topological diversities and heating demands levels. We employ a combined approach using visual inspection and random-walk clustering [29] to select a diverse set of clusters. We assess the diversity of these clusters in terms of both topology (i.e., types and sizes) and cluster-level heating demand powers (in MegaWatts, MW) averaged over time.

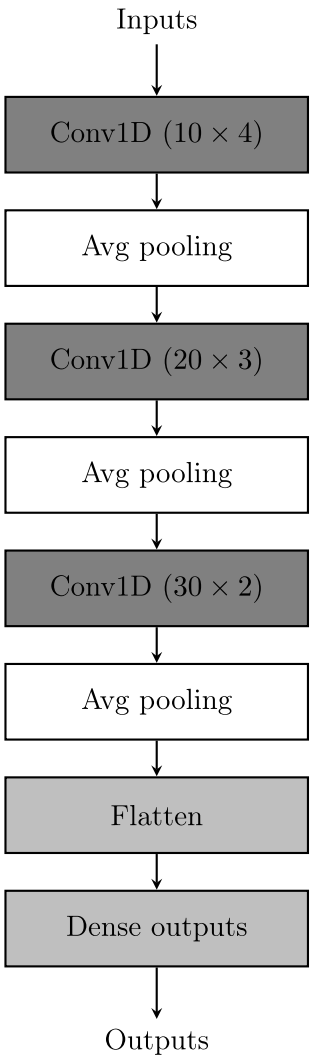
4.2. Hybrid simulation on new scenarios

This topology reduction method aims to significantly reduce computational costs for District Heating Network (DHN) simulations. This allows operators to efficiently test various operational scenarios. Therefore, we assess the proposed hybrid simulation approach validity under two simulation test scenarios, shown in Fig. 6. In first scenario, we consider heating demands profiles from a week in August 2019 (outside the training data), while keeping the generation supply temperature profiles similar to used to generate the training data. In the second scenario, we change the generation supply temperature profiles (constant 90 °C and 100 °C), while considering heating demands from the training data (March 2019).

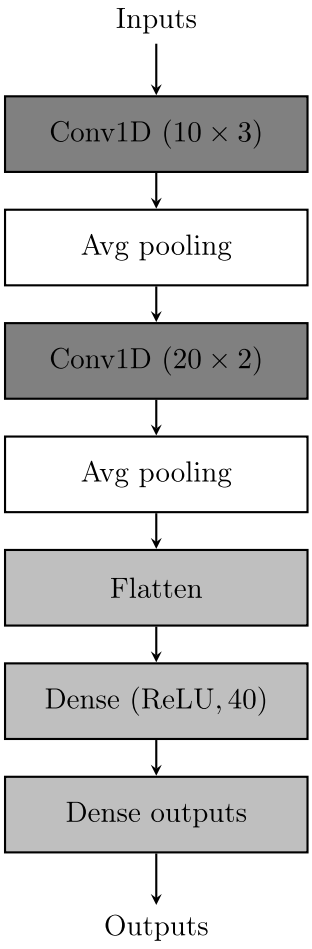
5. Results

5.1. Selected clusters

We select 220 clusters (54 from DHN1, 29 from DHN2, 57 from DHN3 and 80 from DHN4) ensuring significant topological diversity (size and type) and heating demand variations (see Fig. 7). Clusters range from very small (2 consumers) to the largest topologically possible within the networks, as shown in Fig. 7(a). Diversity in external connectivity is reflected by the wide range of cluster types, shown in Fig. 7(b). Also, Fig. 8 shows that the average heating demands of the selected clusters span over a significant range of values. The heating demand of the cluster is simply the sum of the heating demands of its individual consumer.



(a) CNN v1 architecture



(b) CNN v2 architecture

Fig. 4. The considered CNN architectures versions. Number of filters along with their kernel sizes are shown in the 1D convolution layer cases.

Table 2
Our considered ANNs architectures complexities measured in terms of average number of trainable weights and average inference time (in milliseconds ms). We note that presented values are averaged over the selected clusters.

ANNs models	Avg. number of weights	Avg. inference time (ms)
MLP v1	7485	18
MLP v2	26955	21
GRU-NN v1	6635	21
GRU-NN v2	10175	23
CNN v1	**4045**	**18**
CNN v2	8425	21

5.2. Learning performances

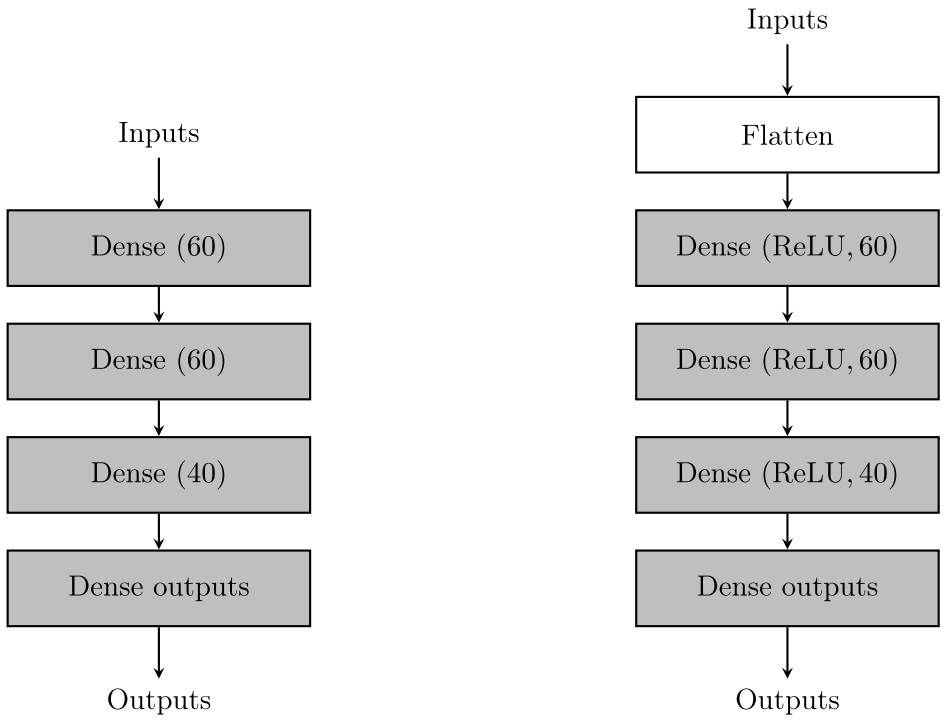
We evaluate the performance of the ANNs models on unseen data (test sets) for each selected cluster. The evaluation metric is the denormalized Mean Absolute Error (MAE) in °C. We compare the distributions of these MAEs to assess the overall performance and generalizability of each ANNs models architectures across the diverse clusters. Since our models have varying complexities, in number of trainable weights and inference times as detailed in Table 2, we assess both accuracy and model complexity when judging the final results.

Table 3
Learning performances of the considered ANNs architectures, averaged over the selected clusters.

ANNs models	Using L1-regularizer	MAEs (°C)
MLP v1	No	0.797(±0.214)
	Yes	0.884(±0.130)
MLP v2	No	0.788(±0.142)
	Yes	0.940(±0.280)
GRU-NN v1	No	0.172(±0.07)
	Yes	0.460(±0.16)
GRU-NN v2	No	0.189(±0.09)
	Yes	0.300(±0.11)
CNN v1	No	0.132(±0.01)
	Yes	0.254(±0.11)
CNN v2	No	**0.115(±0.02)**
	Yes	0.198(±0.11)

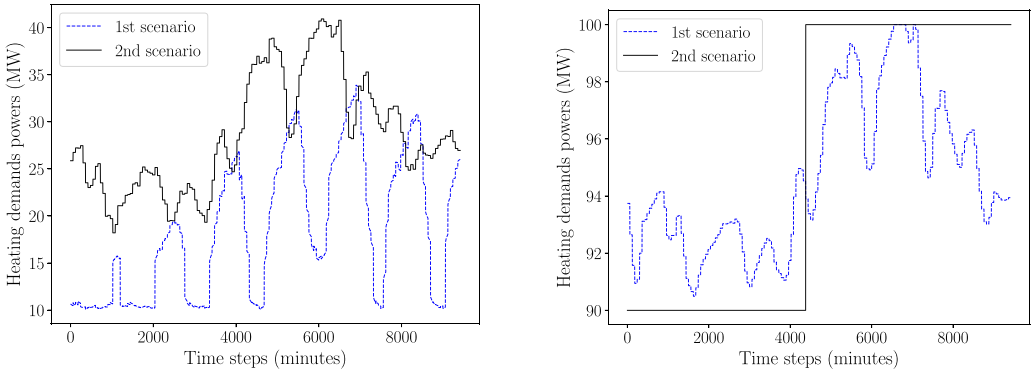
5.2.1. Impact of the use of L1-regularizer

We report mean MAEs with confidence intervals for all ANNs models in Table 3. The results suggest that the use of L1 regularization during training might decrease the model learning performance. To statistically validate this, we employ non-parametric Mann–Whitney



(a) MLP v1 architecture (time-to-time) (b) MLP v2 architecture (flattened sequences)

Fig. 5. The considered MLP architectures versions. Every shown dense layer uses Rectified Linear Unit activation function with shown number of units.



(a) DHN aggregated heating demands (b) Generation supply temperatures of the sources

Fig. 6. Profiles of the aggregated heating demands 6(a) and generation supply temperatures 6(b) for both simulation test scenarios.

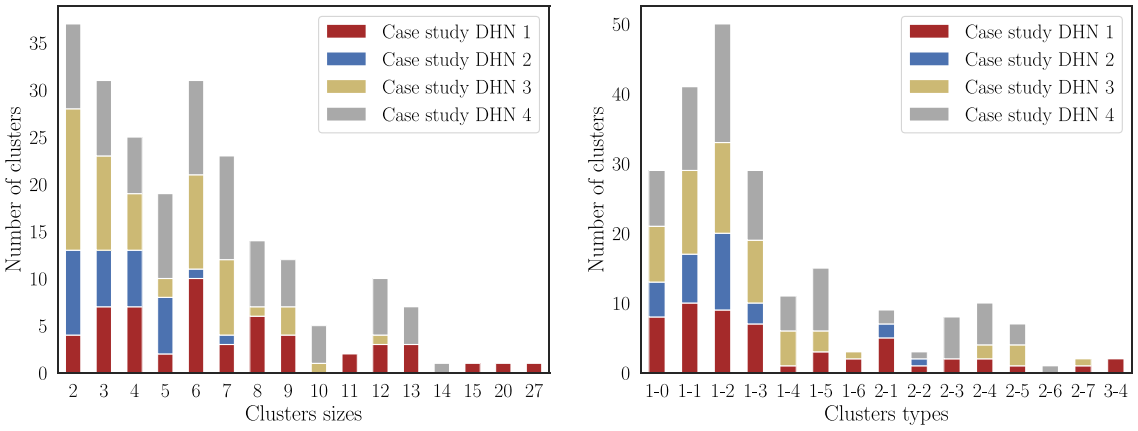
U tests [30] to compare the MAE distributions. The null hypothesis implies that the MAE distributions of the ANNs models are similar with and without L1 regularization. We consider a confidence interval of 95% corresponding to p-value lower than 0.05. All models except MLP v2 (p-value = 0.257) show statistically significant performance degradation with regularization of L1, having p-values lower than 0.05 (MLP v1: $6.03 * 10^{-9}$, GRU-NN v1: $1.76 * 10^{-8}$, GRU-NN v2: $9.83 * 10^{-9}$, CNN v1: $1.31 * 10^{-3}$, CNN v2: $1.9 * 10^{-2}$). This might suggest the potential challenging task of training architecture with high number of weights such as the MLP v2 with or without L1-regularizer. It emphasizes also the use of compact ANNs architectures over large architectures.

5.2.2. ANNs models comparison

Table 3 suggests that MLP architectures which exhibit the highest MAEs are less suitable for capturing cluster underlying physics compared to GRU-NN and CNN models, shown in Fig. 9. We further validate this result statistically using the Mann–Whitney U tests. Obtained p-values for different tests are reported in Table 4. In mean values, MLP v1 performs slightly better than v2, which is validated with the test result that produces the p-value < 0.05. However, both are excluded due to overall poor performances (Table 3). GRU-NN v1 and GRU-NN v2 do not show statistically significant differences (p-value > 0.05), suggesting that v1 performs similarly to v2 despite having fewer weights and faster inference. However, CNN v2 outperforms CNN v1 (p-value < 0.05), justifying its increased complexity. Finally, CNN v2 and GRU-NN v1 show no statistically significant differences (p-value > 0.05). This suggests both models to be equally effective in learning the clusters physics, with GRU-NN v1 offering an advantage in terms of computational complexity as shown in Table 2.

5.2.3. Analysis on the impacts of the clusters topology and heating demands

We further analyze how cluster properties affect GRU-NN v1 and CNN v2 performance (best identified models, see previous section).



(a) Selected clusters grouped by their sizes (number of consumer nodes)

(b) Selected clusters grouped by their types (number and directions of external edges)

Fig. 7. Distributions of the selected clusters topology in function of their respective sizes 7(b) and types 7(b). We also show the number of selected clusters from the four different synthetic District Heating Networks.

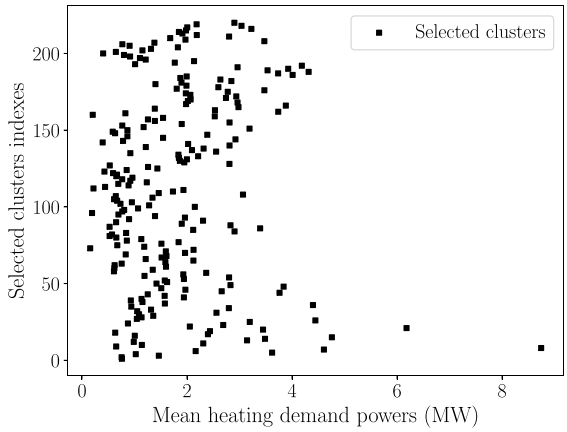


Fig. 8. Selected clusters average heating demand powers.

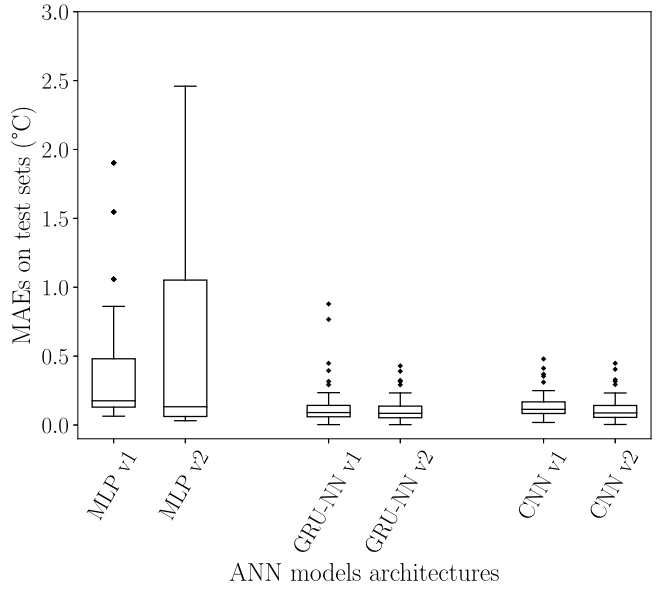


Fig. 9. Distributions of the MAEs performances of the considered ANNs architectures without the use of L1-regularizer.

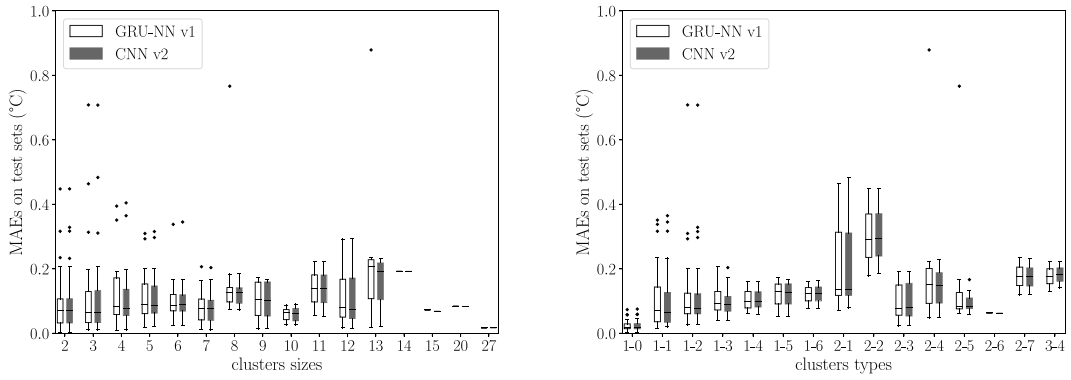
Table 4
P-values of different statistical test using Mann–Whitney U-test for different distribution MAEs comparisons. Null hypothesis infers no statistical differences between the distributions.

ANNs differences compared	*p*-value
MLP v1 and MLP v2	$1 * 10^{-3}$
GRU-NN v1 and GRU-NN v2	0.578
CNN v1 and CNN v2	$6.98 * 10^{-4}$
MLP v1 and GRU-NN v1	$7.11 * 10^{-27}$
MLP v1 and CNN v2	$2.77 * 10^{-23}$
GRU-NN v1 and CNN v2	0.695

- Cluster size: Fig. 10(a) shows similar model performance across cluster sizes, further supported by a weak negative correlation coefficient (−0.05). It implies that the sizes of the clusters do not impact directly the ANNs models ability to learn the underlying physics of the clusters. Besides, large clusters may potentially provide larger computational cost reduction than small clusters.
- Cluster type: In general, the ANNs models perform significantly better on type 1–0 clusters (Fig. 10(b)), due to the lower number of temperatures profiles to learn and predict. However, performance varies within the same type. On one hand, this result suggests that our framework can handle diverse external connectivity (*i.e.*, types). On the other hand, further investigation on the link between performances and the clusters topology metrics might be beneficial, although outside of the scope of this study.
- Cluster heating demands: The amplitude of heat demands appears to have minimal influence on the learning performance of the ANNs model (Fig. 11). The MAEs distributions do not show a clear correlation with the average cluster heating demand (MW). Low Pearson correlation coefficients (0.03 for GRU-NN v1 and 0.02 for CNN v2) support this observation. This suggests that our framework can handle clusters with wide ranges of heating demands levels, allowing for the replacement of low- or high-consumption clusters without significant performance changes.

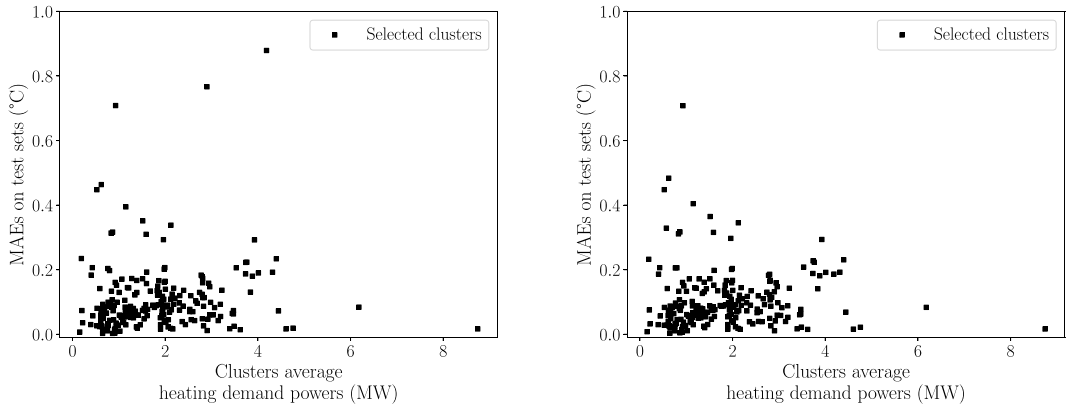
5.3. Hybrid simulation

We evaluate the hybrid simulation using case study DHN1. We consider eight clusters, five of type 1–0 (K1–K5) and three of type 1–1 (K6–K8) (Fig. 12). Type 1–0 clusters are chosen due to their superior learning performance (Section 5.2.3). However, type 1–1 clusters are included to gauge the impact of cluster substitutions on the supply side. The average mass flow rates of the water incoming to the clusters are



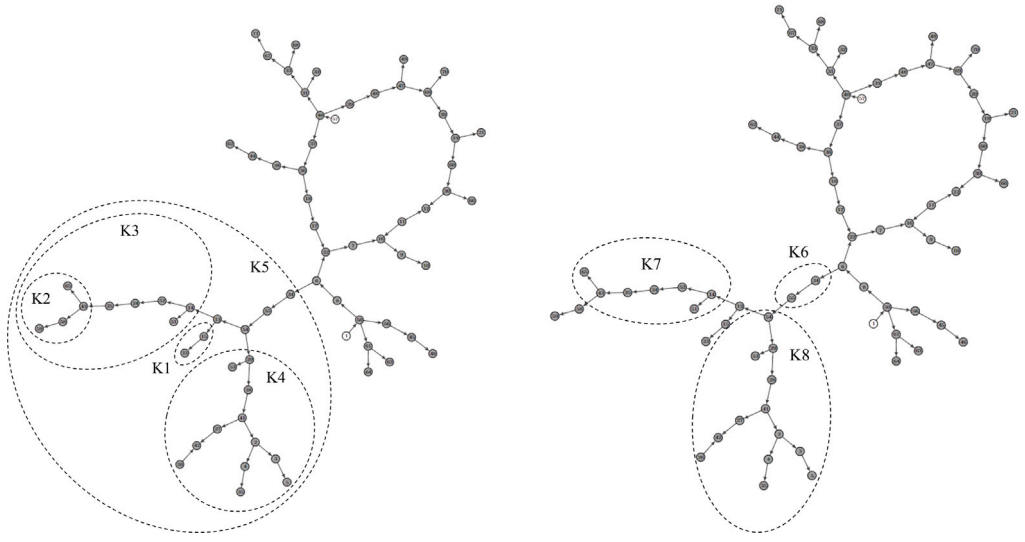
(a) MAEs in function of clusters sizes (b) MAEs in function of clusters types

Fig. 10. Distribution of the models GRU-NN v1 and CNN v2 learning performances (MAEs) over the selected clusters, with respect to their sizes and types.



(a) MAEs of the GRU-NN v1 (b) MAEs of the CNN v2

Fig. 11. Distribution of the models GRU-NN v1 and CNN v2 learning performances (MAEs) over the selected clusters, with their averaged heating demands.



(a) Type 1-0 considered clusters (b) Type 1-1 considered clusters

Fig. 12. Case study DHN1 with eight clusters (delimited in dashed lines) considered to validate the hybrid simulation framework.

Table 5
Mass flow rates of the water (kg/s) entering the clusters, averaged over time, on both simulation scenarios.

Clusters	K1	K2	K3	K4	K5	K6	K7	K8
1st scenario	3.6	8.6	15.3	24.3	49.9	49.9	19.9	45.0
2nd scenario	5.1	12.2	21.9	36.4	73.2	73.2	28.4	66.0

Table 6
Accuracy performances of the hybrid simulations on the case study DHN1 by substituting the mentioned clusters with their corresponding twin nodes, measured by physics-based metrics.

Clusters	K1	K2	K3	K4	K5	K6	K7	K8
sizes	2	4	9	12	27	2	8	13
Thermal loads (%)								
1st scenario								
GRU-NN v1	0.043	0.119	0.087	0.054	0.115	0.084	0.294	0.093
CNN v2	0.045	0.125	0.085	0.057	0.122	0.086	0.326	0.110
2nd scenario								
GRU-NN v1	0.007	0.016	0.024	0.016	0.031	0.033	0.161	0.038
CNN v2	0.007	0.018	0.023	0.016	0.027	0.037	0.149	0.044
Produced heat powers (%)								
1st scenario								
GRU-NN v1	0.012	0.014	0.025	0.035	0.088	0.030	0.033	0.039
CNN v2	0.013	0.015	0.024	0.037	0.093	0.033	0.036	0.053
2nd scenario								
GRU-NN v1	0.002	0.002	0.007	0.011	0.024	0.021	0.019	0.021
CNN v2	0.002	0.002	0.006	0.010	0.020	0.021	0.018	0.018

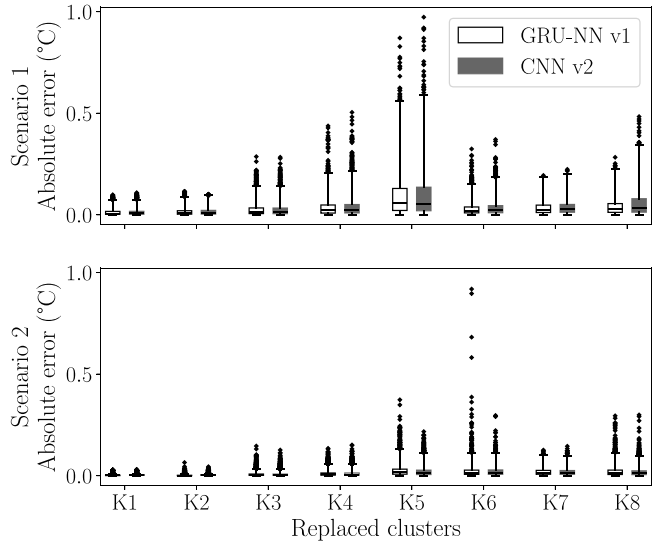


Fig. 13. Distribution of errors by the hybrid simulations on preserving the temperatures of the return water at the sources, measured in absolute differences.

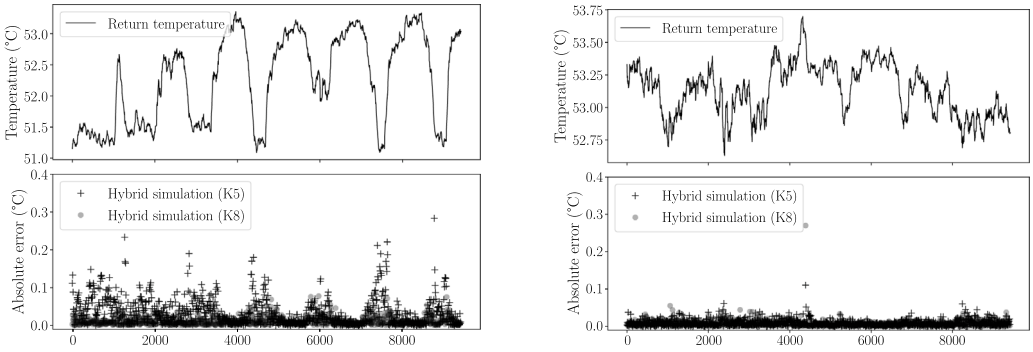
Table 7
Mean relative errors (%) on conserving the thermal loads (upstream | downstream) of the type 1-1 considered clusters.

Clusters	K6	K7	K8
1st scenario			
GRU-NN v1	0.038 | 0.130	0.051 | 0.538	0.058 | 0.129
CNN v2	0.043 | 0.130	0.056 | 0.596	0.078 | 0.142
2nd scenario			
GRU-NN v1	0.027 | 0.039	0.029 | 0.293	0.030 | 0.046
CNN v2	0.027 | 0.048	0.149 | 0.271	0.044 | 0.060

presented in Table 5. These rates correspond to both supply entering and return leaving the clusters. The entering mass flow rates reflect the heat consumption and losses within the clusters, leading to a positive correlation with cluster size for type 1–0 clusters. However, for type 1–1 clusters, additional downstream thermal requirements also influence the entering mass flow rates.

5.3.1. Accuracy performances

Table 6 summarizes the accuracy performances of the hybrid simulation results. Consistent with prior findings, GRU-NN v1 and CNN v2 exhibit similar performance. Both models effectively preserve the temperatures of outgoing water from replaced clusters in both scenarios, reflected by low errors on the thermal loads. Indeed, lower errors on predicting the temperatures of outgoing water result in lower errors on the heat power at the interfaces of the clusters. These performances do not directly correlate with size of the clusters. Instead, they reflect the surrogate ANNs models capability to capture the underlying physics, supporting results in Section 5.2.3. Type 1–1 clusters have two interfaces with the rest of the network (upstream and downstream). Reported values in Table 6 represent the averaged errors on the thermal loads at both interfaces which are individually reported in Table 7. In contrast, the errors on the produced heat powers by the sources depend on the propagation of errors in the leaving return temperatures predictions, from the twin nodes back to the sources. Results in Tables 6 and 7 show that this error propagation is influenced by two contributions: the level of upstream thermal load, linked to the mass flow rate of the incoming water flows of the clusters, and errors on this upstream load conservation, reflecting the errors on predicting outgoing return temperatures. For type 1–1 cluster substitutions, an additional error comes from the predictions of the outgoing supply temperatures. This error can influence downstream substation consumption and consequently, their return temperatures. Generally, we observe higher error values and distributions under scenario 1 compared to scenario 2 (Fig. 13). This is consistent with the greater variability of return temperatures observed in scenario 1 compared to scenario 2 (Fig. 14). In contrast,



(a) Under the 1^{st} scenario (b) Under the 2^{nd} scenario

Fig. 14. Profiles of the return temperatures at the source under both simulation scenarios using physical simulation and absolute errors from the hybrid simulations.

scenario 2 exhibits a more abrupt change in generation supply temperatures. However, according the errors distributions shown in Fig. 13, such supply temperature profile change does not have major impact on the clusters substitution performances, regardless the types (1–0 and 1–1). Although, we observe some wide outliers performances from the replacement of cluster K6 which can be explained by its proximity to the heat sources (see Fig. 12).

5.3.2. CPU time reduction

Beyond accuracy, we investigated the potential of hybrid simulations to reduce computational time. Since both architectures have similar inference times (Table 2), the CPU time for hybrid simulations is independent of the chosen model (GRU-NN v1 or CNN v2). Therefore, we focus on the CPU time of hybrid simulations using GRU-NN v1 for comparison with the physical simulation (Fig. 15). This figure demonstrates that using larger clusters leads to significant computational gains. Smaller clusters substitutions do not compensate for the ANNs model inference times with their reduced number of physical equations. Consequently, our largest cluster (K5, 27 nodes, 39% of DHN1 nodes) reduces CPU time by 27% (Fig. 15) while maintaining sources heat power mean relative errors below 0.1%. While this simulation time reduction might appear modest, our results reveal key factors crucial for evaluating the profitability of this approach in real-world operation. First, cluster learning, which takes roughly 30 min, is a one-time investment (*e.g.*, first implementation year). The true benefit lies in the cumulative time saved through repeated predictions during the multiple simulations needed for operational optimization. Second, as shown in Fig. 15, computational gains become more significant as cluster size increases. The strength of our approach lies in its ability to effectively reduce simulation time for large District Heating Networks (DHNs) by leveraging large type 1–0 clusters, all while maintaining high accuracy.

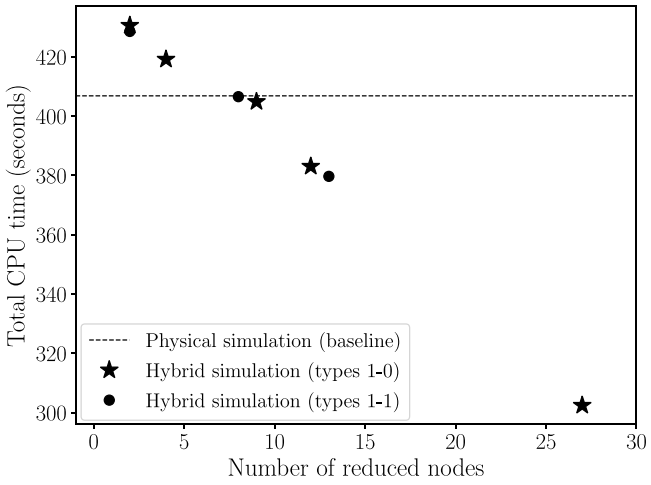


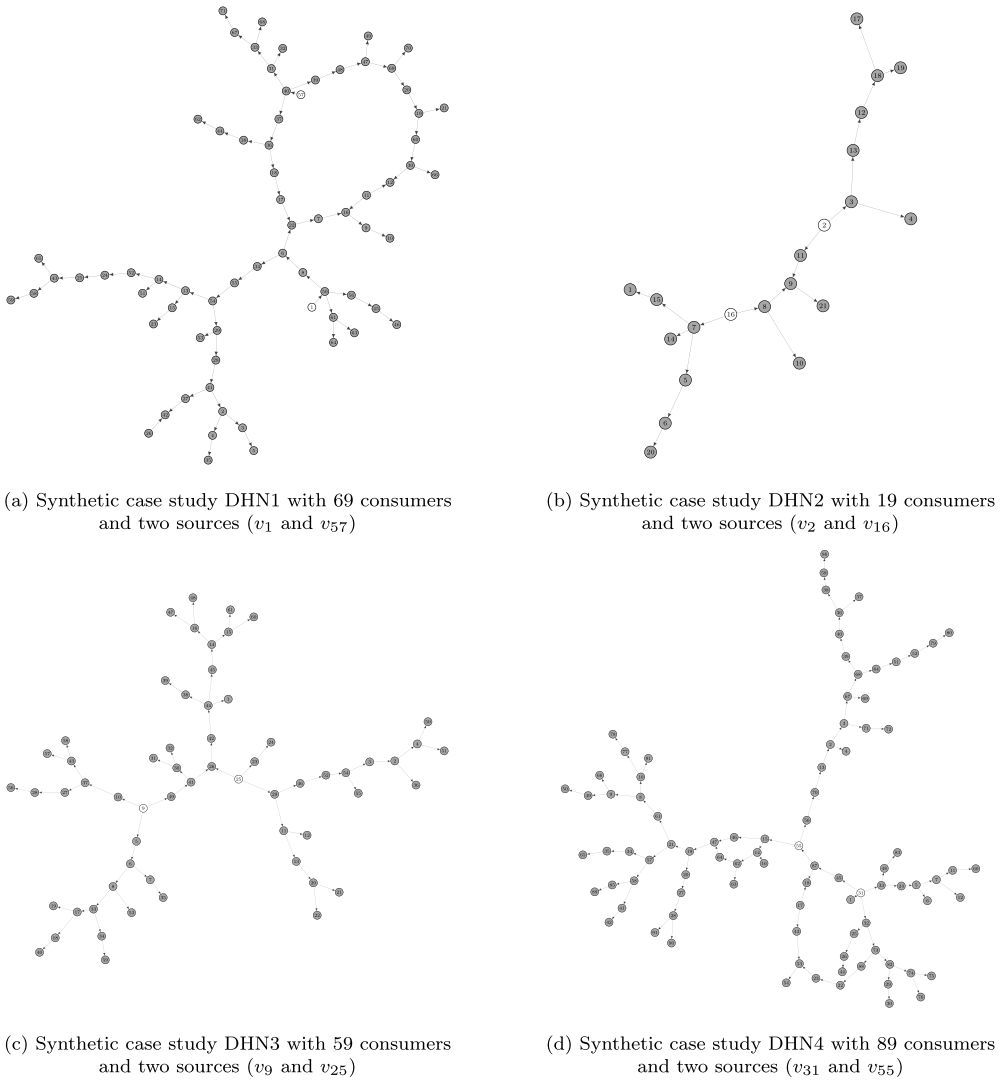
Fig. 15. Comparison of the CPU time required by the hybrid simulations on different reduced networks, using GRU-NN v1, and by physical simulation on original network.

Table A.1
Reference U-factors used. Each category contains approximately 33% of the substations of the networks.

Substation categories	U-factor value [$W m^{-2} K^{-1}$]
Ancient buildings	3.4
Recent buildings	0.84
Offices	2.5

6. Discussion and conclusion

This study investigated an aggregation framework using Artificial Neural Networks (ANNs) to reduce the computational costs of simulating District Heating Networks (DHNs). By replacing user-driven selected consumer clusters with surrogate ANNs models, the framework simplifies the DHNs topology. Our experiments on four case study DHNs demonstrate the ability of the ANNs models to learn the physical dynamics of diverse clusters, regardless of topology or heating demand levels. In particular, external connectivity appears to have a greater impact on learning performance than cluster size or heating demand level. We further coupled the predictions of the ANNs models with physical simulations, achieving significant reductions in simulation time while maintaining accurate source power conservation. More precisely, these reduction of simulation times correlate directly with the size of the replaced clusters. However, the observed performance gap between clusters indicates factors beyond type, size, and heating demand are affecting the ability of the surrogate ANNs models to learn the underlying physical dynamics of the clusters. To understand these influences better, future research will explore additional metrics and assess their impact on the learning performances. This could lead to more relevant cluster identifications for the aggregation framework. Furthermore, identifying the most influential features within the input features of the ANNs models may allow for a reduction in computational cost by using fewer features. Future work would delve into this possibility as well.



(a) Synthetic case study DHN1 with 69 consumers and two sources (v_1 and v_{57})

(b) Synthetic case study DHN2 with 19 consumers and two sources (v_2 and v_{16})

(c) Synthetic case study DHN3 with 59 consumers and two sources (v_9 and v_{25})

(d) Synthetic case study DHN4 with 89 consumers and two sources (v_{31} and v_{55})

Fig. B.1. Generated synthetic District Heating Networks (DHN) which serve as our case study networks. Consumers are in gray-filled nodes and sources are white nodes.

CRediT authorship contribution statement

Dubon Rodrigue: Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Formal analysis, Conceptualization. **Mohamed Tahar Mabrouk:** Writing – review & editing, Supervision, Software, Formal analysis. **Bastien Pasdeloup:** Writing – review & editing, Supervision, Software, Formal analysis. **Patrick Meyer:** Writing – review & editing, Software, Resources, Formal analysis. **Bruno Lacarrière:** Writing – review & editing, Software, Resources, Formal analysis.

Declaration of competing interest

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

Data availability

Case study District Heating Networks topology files, the list of selected clusters, hybrid simulation results, training codes, data pipeline of the machine learning models and performances results are made publicly available on our GitHub repository (https://github.com/drod-96/smart_clusters_v1.git). Due to size limitations of GitHub, simulation results and forecast heating demands files have been uploaded and published on Mendeley Data repository (https://data.mendeley.com/datasets/77stj44drm/1). The physical simulation codes used in this work will be made publicly available in the future, alongside a dedicated paper that delves deeper into their functionalities.

Acknowledgments

We would like to thank the CNRS for granting this work access to the HPC/AI resources of the “Institut du développement et des ressources en informatique scientifique” under the allocation 20XX-AD011014376, made by GENCI.

Appendix A. Heating demands and generation supply temperatures models

To mimic realistic heating demands of buildings and offices, we generate values based on an adapted heating law, governed by Eq. (A.1). This law relates heating demands to outdoor temperatures $\boldsymbol{T}_{\text{outdoor}}$. A threshold temperature $T_{\text{threshold}}$ of 18 °C is fixed. Domestic Hot Water (DHW) demands are uniformly distributed between 100 kW and 200 kW. Three U-factors (heat exchange rates) are considered, representing ancient buildings, recent buildings and offices, presented in Table A.1.

Each consumer node is uniformly and randomly assigned to one of these categories. The total heat exchanger surface area is denoted by S. We use publicly available outdoor temperature data for Nantes city center for the entire year 2019, obtained from the NASA Power website [28]. During hot hours with $\boldsymbol{T}_{\text{outdoor}}(t) \geq T_{\text{threshold}}$ for t denoting the hour, heating demands are assumed to be 20% of the peak demand, obtained at $T_{\text{outdoor}} = -6°\text{C}$ (*e.g.*, hot water domestic needs). Similarly, generation supply temperatures from the sources depend inversely on the outdoor temperatures which is governed by Eq. (A.2). In this work, the generation supply temperatures are bounded between 90°C and 100°C.

$$\forall (v \in \mathcal{V}_C, t),\ \boldsymbol{D}_v(t) = \text{DHW} + \begin{cases} \text{U} \cdot S \cdot \left(\boldsymbol{T}_{\text{outdoor}}(t) - T_{\text{threshold}}\right) & \text{if } \boldsymbol{T}_{\text{outdoor}}(t) < T_{\text{threshold}} \\ 24 \cdot 0.2 \cdot \text{U} \cdot S & \text{if } \boldsymbol{T}_{\text{outdoor}}(t) \geq T_{\text{threshold}} \end{cases} \tag{A.1}$$

$$\forall (p,t),\quad \boldsymbol{T}_{ss_p}(t) = \begin{cases} 100\ °\text{C} & \text{if } \boldsymbol{T}_{outdoor}(t) \leq -1°\text{C} \\ a \cdot \boldsymbol{T}_{outdoor}(t) - b & \text{if } \boldsymbol{T}_{outdoor}(t) \in]-1, 15[\ °\text{C} \\ 90\ °\text{C} & \text{if } \boldsymbol{T}_{outdoor}(t) \geq 15°\text{C} \end{cases} \tag{A.2}$$

Appendix B. Synthetic district heating networks considered

See Fig. B.1.

References

[1] Lund H, Möller B, Mathiesen BV, Dyrelund A. The role of district heating in future renewable energy systems. Energy 2010;35(3):1381–90. http://dx.doi.org/10.1016/j.energy.2009.11.023.

[2] Mans M, Blacha T, Remmen P, Müller D. Automated model generation and simplification for district heating and cooling networks. In: Linköping electronic conference proceedings, vol. 157. Regensburg, Germany; 2019, p. 179–86. http://dx.doi.org/10.3384/ecp19157179, URL: https://ep.liu.se/en/conference-article.aspx?series=ecp&issue=157&Article_No=18.

[3] Ayele G, Mabrouk M, Haurant P, Laumert B, Lacarrière B. Pseudo-dynamic simulation on a district energy system made of coupling technologies. In: Proceedings of ECOS 2018. Guimaraes, Portugal: University of Minho; 2018, URL: https://hal.science/hal-01879566.

[4] Cosentino S, Elisa G, Melli R, Sciacovelli A, Enrico S, Toro C, et al. Identification of the optimal operational strategy of a large district heating network through POD modeling. In: Proceedings of ECOS 2014. Turku, Finland: Abo Akademi University; 2014, URL: https://www.academia.edu/27636427/Identification_of_the_optimal_operational_strategy_of_a_large_district_heating_network_through_POD_modeling.

[5] Rein M, Mohring J, Damm T, Klar A. Reduction of a district heating model using network decomposition. Proc Appl Math Mech 2019;19(1):e201900038. http://dx.doi.org/10.1002/pamm.201900038, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/pamm.201900038.

[6] Rein M, Mohring J, Damm T, Klar A. Optimal control of district heating networks using a reduced order model. Optim Control Appl Methods 2020;41(4):1352–70. http://dx.doi.org/10.1002/oca.2610, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/oca.2610.

[7] Simonsson J, Atta KT, Birk W. Reduced-order modeling of thermal dynamics in district energy networks using spectral clustering. 2022, http://dx.doi.org/10.48550/arXiv.2202.09259, arXiv, URL: http://arxiv.org/abs/2202.09259, arXiv:2202.09259 [cs, eess].

[8] Biener W, Garcia Rosas KR. Grid reduction for energy system analysis. Electr Power Syst Res 2020;185:106349. http://dx.doi.org/10.1016/j.epsr.2020.106349.

[9] Loewen A, Wigbels M, Althaus W, Augusiak A, Renski A. Structural simplification of complex DH-networks. Euroheat Power Fernwaerme Int 2001;30:42–4.

[10] Larsen HV, Pálsson H, Bøhm B, Ravn HF. Aggregated dynamic simulation model of district heating networks. Energy Convers Manage 2002;43(8). http://dx.doi.org/10.1016/S0196-8904(01)00093-0.

[11] Larsen V H, Bøhm B, Wigbels M. A comparison of aggregated models for simulation and operational optimisation of district heating networks. Energy Convers Manage 2004;45(7–8):1119–39. http://dx.doi.org/10.1016/j.enconman.2003.08.006.

[12] Kane M, Rolle J. "Quantum networks": A new approach for representing a network and evaluating hydraulic and thermal losses in district heating/cooling systems. In: Proceedings of ECOS 2020. Osaka, Japan; 2020, URL: https://arodes.hes-so.ch/record/7042.

[13] Franki V, Majnarić D, Višković A. A comprehensive review of artificial intelligence (AI) companies in the power sector. Energies 2023;16(3):1077. http://dx.doi.org/10.3390/en16031077, Number: 3 Publisher: Multidisciplinary Digital Publishing Institute.

[14] Gianniou P, Liu X, Heller A, Nielsen PS, Rode C. Clustering-based analysis for residential district heating data. Energy Convers Manage 2018;165:840–50. http://dx.doi.org/10.1016/j.enconman.2018.03.015.

[15] Calikus E, Nowaczyk S, Sant'Anna A, Gadd H, Werner S. A data-driven approach for discovering heat load patterns in district heating. Appl Energy 2019;252:113409. http://dx.doi.org/10.1016/j.apenergy.2019.113409.

[16] Huang B, Wang J. Applications of physics-informed neural networks in power systems — A review. IEEE Trans Power Syst 2023;38(1):572–88. http://dx.doi.org/10.1109/TPWRS.2022.3162473, Conference Name: IEEE Transactions on Power Systems.

[17] Paudel S, Elmtiri M, Kling WL, Corre OL, Lacarrière B. Pseudo dynamic transitional modeling of building heating energy demand using artificial neural network. Energy Build 2014;70:81–93. http://dx.doi.org/10.1016/j.enbuild.2013.11.051.

[18] Adam A, Fang H, Li X. Effective thermal conductivity estimation using a convolutional neural network and its application in topology optimization. Energy AI 2024;15:100310. http://dx.doi.org/10.1016/j.egyai.2023.100310.

[19] Petrucci A, Barone G, Buonomano A, Athienitis A. Modelling of a multi-stage energy management control routine for energy demand forecasting, flexibility, and optimization of smart communities using a recurrent neural network. Energy Convers Manage 2022;268:115995. http://dx.doi.org/10.1016/j.enconman.2022.115995.

[20] Owerko D, Gama F, Ribeiro A. Optimal power flow using graph neural networks. 2019, arXiv, URL: http://arxiv.org/abs/1910.09658, arXiv:1910.09658 [cs, eess].

[21] Benonysson A, Bøhm B, Ravn HF. Operational optimization in a district heating system. Energy Convers Manage 1995;36(5). http://dx.doi.org/10.1016/0196-8904(95)98895-T.

[22] Ayele GT, Mabrouk MT, Haurant P, Laumert B, Lacarrière B. Optimal placement and sizing of heat pumps and heat only boilers in a coupled electricity and heating networks. Energy 2019;182. http://dx.doi.org/10.1016/j.energy.2019.06.018.

[23] Ayele GT, Haurant P, Laumert B, Lacarrière B. An extended energy hub approach for load flow analysis of highly coupled district energy networks: Illustration with electricity and heating. Appl Energy 2018;212:850–67. http://dx.doi.org/10.1016/j.apenergy.2017.12.090.

[24] Hägg R. Dynamic simulation of district heating networks in dymola [Ph.D. thesis], Lund University; 2016, URL: https://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=8883842&fileOId=8883845.

[25] Chung J, Gulcehre C, Cho K, Bengio Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. 2014, http://dx.doi.org/10.48550/arXiv.1412.3555, arXiv, URL: http://arxiv.org/abs/1412.3555, arXiv:1412.3555 [cs].

[26] Kiranyaz S, Avci O, Abdeljaber O, Ince T, Gabbouj M, Inman DJ. 1D convolutional neural networks and applications: A survey. Mech Syst Signal Process 2021;151:107398. http://dx.doi.org/10.1016/j.ymssp.2020.107398.

[27] Klambauer G, Unterthiner T, Mayr A, Hochreiter S. Self-normalizing neural networks. 2017, http://dx.doi.org/10.48550/arXiv.1706.02515, arXiv, URL: http://arxiv.org/abs/1706.02515, cs,stat version: 5.

[28] POWER | Data access viewer, 2023. URL: https://power.larc.nasa.gov/data-access-viewer/.

[29] Pons P, Latapy M. Computing communities in large networks using random walks (long version). In: Computer and information sciences, vol. 3733. Berlin: Springer; 2005, http://dx.doi.org/10.1007/11569596_31, URL: https://link.springer.com/chapter/10.1007/11569596_31#citeas, arXiv:physics/0512106.

[30] Mann HB, Whitney DR. On a test of whether one of two random variables is stochastically larger than the other. Ann Math Stat 1947;18(1):50–60. http://dx.doi.org/10.1214/aoms/1177730491, Publisher: Institute of Mathematical Statistics.