Energy 350 (2026) 140746

Contents lists available at ScienceDirect

Energy

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

Distributed model predictive control for district heating networks considering building and network flexibility

Felix Langner *, Jörg Matthes, Veit Hagenmeyer

Karlsruhe Institute of Technology, Institute for Automation and Applied Informatics, Eggenstein-Leopoldshafen, Germany

ARTICLE INFO

Keywords:
Demand response
Distributed control
District heating networks
Model predictive control
Residential buildings

ABSTRACT

The grid-responsive operation of large-scale heat pumps in District Heating Networks (DHNs) requires advanced control strategies to coordinate the DHNs' inherent thermal storage with power grid demands. The present work introduces a Distributed Model Predictive Control (DMPC) that leverages both building-level and network-level flexibility to adjust heat pump operation in response to dynamic pricing. The DMPC is based on the Alternating Direction Method of Multipliers algorithm to avoid privacy-critical information flows and improve scalability compared to standard centralized MPC (CMPC). To achieve this, each building solves a local control problem and exchanges only non-sensitive information with the DHN operator. The DMPC is comprehensively compared against a CMPC on a literature-based benchmark DHN across nine one-week simulation scenarios varying in price and weather conditions. The DMPC successfully achieves control performance comparable to the CMPC with excess costs below 1.5% while maintaining thermal comfort and avoiding privacy-critical information flow. Additionally, the computation time is reduced by 15.5% to 74.3% compared to the CMPC, highlighting the improved scalability. The computation time savings substantially depend on the number of iterations necessary for the DMPC to converge, emphasizing the importance of avoiding excess iterations. These results highlight the potential of distributed control to facilitate scalable, privacy-preserving, and efficient demand response in future DHNs.

1. Introduction

A core strategy for decarbonizing the heating sector is the extension of District Heating Networks (DHNs) [1]. The "Heat Roadmap Europe" studies estimate that by 2050, district heating could provide 50% of the total heat demand across the EU27 countries, with large-scale heat pumps accounting for 25%–30% of this supply [2]. As a result, DHNs offer significant potential to provide demand flexibility for the power grid. This flexibility can be harnessed through Demand Response (DR) programs, which incentivize District Heating Operators (DHOs) to operate heat pumps grid-interactively by adjusting heating in response to dynamic electricity prices. For DHNs equipped with large-scale heat pumps to participate in a DR program, heat flexibility is required to align the heat consumption with the heat pump's operation.

In general, three sources of heat flexibility exist within a DHN [3]: the heat carrier, thermal storage tanks, and the thermal mass of the buildings connected to the DHN. Particularly, the building flexibility has been found to be the dominating source of flexibility [4]. While classical control of DHNs aims to balance heat supply and demand, advanced control can purposefully utilize the previously mentioned sources of flexibility to create unequal supply and demand, thereby reducing costs and environmental impact [5]. Advanced controls for DHNs are typically based on centralized optimization, such as Centralized Model Predictive Control (CMPC) [5]. In CMPC, a single centralized controller determines the control actions for all consumers in the DHN. While effective, a centralized controller lacks scalability [6] and requires the gathering of large amounts of data centrally, including privacy-critical information such as the occupancy of buildings connected to the DHN [7]. Due to these privacy concerns, a centralized controller may be unable to leverage the buildings' flexibility, thereby limiting the DHN's ability to respond to grid signals. In contrast, distributed control for DHNs has been found vital for reducing privacy-critical information flows and improving scalability [5], thereby enabling the integration of building flexibility.

Fig. 1 illustrates the architectures for centralized and distributed optimization.

In the centralized configuration, a single optimizer collects all data of the DHN and buildings, including sensitive measurements, system models, and occupancy information. As the number of buildings increases, the size of the optimization problem grows.

* Corresponding author.
E-mail address: felix.langner@kit.edu (F. Langner).

https://doi.org/10.1016/j.energy.2026.140746
Received 26 August 2025; Received in revised form 7 January 2026; Accepted 12 March 2026
Available online 13 March 2026

Nomenclature

Acronyms

ADMM	Alternating Direction Method of Multipliers
CF-VT	Constant Flow Variable Temperature
CMPC	Centralized Model Predictive Control
DHN	District Heating Network
DHO	District Heating Operator
DMPC	Distributed Model Predictive Control
DR	Demand Response
VF-CT	Variable Flow Constant Temperature
VF-VT	Variable Flow Variable Temperature

Parameters

Δp	Pressure drop in Pa
Δt	Sample time in s
$\dot{m}$	Water mass flow in kg/s
γ	Penalty parameter for ADMM
$\overline{T}$	Upper temperature bound in °C
ρ	Water density in kg/m^3
$\underline{T}$	Lower temperature bound in °C
C	Heat capacity in J/K
c_p	Water specific heat capacity in J/(kg K)
f_{sol}	Solar heat gain factor
N	Number of time steps in the prediction horizon
R	Thermal resistance in K/W
v	Water velocity in m/s

Indices

i	Iteration index
j	Building index
k	Time index

Variables

λ	Dual variable
ϕ	Heat flow in W
p_{buy}	Electricity price in EUR/MWh
P_{hp}	Heat pump power consumption in W
T_{amb}	Ambient air temperature in °C
T_{air}	Indoor air temperature in °C
T_r	Return temperature in °C
T_s	Supply temperature in °C
T_w	Wall temperature in °C

In contrast, the distributed configuration allows each building and the DHN to conduct their own local optimizations. This approach ensures that sensitive data remains local, while only non-sensitive information is shared. In the distributed configuration, buildings can perform their optimizations in parallel, improving scalability compared to the centralized configuration.

While distributed control for DHNs has been found crucial for enabling the integration of building flexibility [5], research on this topic remains scarce [8].

1.1. Related work

Cai et al. [7] propose a distributed coordination of buildings' heat demands for DR while considering a global constraint of the pumping power. By employing the Alternating Direction Methods of Multipliers (ADMM), their distributed optimization achieves identical results to the centralized one, while preserving information privacy. Similarly, Qin et al. [9] propose an ADMM-based distributed control for the cost-minimal operation of a fifth-generation district heating and cooling network. Compared to the centralized control, the distributed control incurs between 0.3% and 5% excess costs across two network sizes and three price signals. Taylor et al. [10] apply ADMM to decompose a large optimization problem for the DR-based control of an integrated electricity grid and DHN into smaller optimization problems. This alleviates the computational load while achieving comparable results to the centralized optimization. In contrast to the previously mentioned works, Maurer et al. [8] utilize Optimality Condition Decomposition to maximize social welfare in a DHN distributedly. While the approach achieves satisfactory convergence for a small network, convergence issues arise for larger networks.

All of the studies mentioned above neglect the temperature dynamics of either the DHN [7,9,10] or of the buildings [8,10]. As a result, either the flexibility potential of the heat carrier or the buildings' thermal mass cannot be leveraged, limiting the DHN's capability for DR. In contrast, studies that consider both the building and network flexibility [11,12] rely on centralized optimization, requiring privacy-critical information flows as highlighted in Table 1. Consequently, the present study bridges the research gap between privacy-preserving distributed optimization[1] and energy flexibility in DHNs with the following main contributions:

1. The development of a distributed Model Predictive Control (DMPC) method for the DR of DHNs to preserve privacy, foster scalability, and reduce computational requirements.
2. Both the temperature dynamics of the buildings connected to the DHN and of the heat carrier are modeled. This enables the DMPC to utilize the flexibility of the buildings' thermal mass and of the heat carrier in the DHN to increase the DHN's capacity for DR.

1.2. Structure of the present paper

The remainder of this paper is organized as follows. Section 2 introduces the considered DHN, the DMPC based on ADMM, and the case study for the DMPC's evaluation. The results are presented and discussed in Section 3. Finally, the paper is concluded in Section 4 with a discussion and an outlook for future work.

2. Methods

2.1. The AROMA network

The considered DHN is called AROMA and is widely used in literature as a benchmark DHN to develop and evaluate control algorithms (e.g., [15–18]). While there are various modifications of the network that include multiple producers [17] or additional storage [17,18], the present work focuses on the original network configuration as proposed by Krug et al. [15]. The utilized AROMA network is visualized in Fig. 2.

The network comprises one heat producer, which in this work is assumed to be a large-scale heat pump and has a total pipe length of 7262.4 m. Further, the network includes five large-scale consumers, representing aggregated thermal demands in the range of multiple hundred kilowatts. Since the focus of the present work is the provision of flexibility of individual buildings, each of these consumers is modeled

[1] There is no consensus about the nomenclature of control architectures. We will use "distributed" in the general sense of [13] as meaning that participants exchange information to solve a coupled, global optimization problem. Other authors refer to the architecture of our approach as "hierarchical-distributed" [14] or "hybrid" [5].

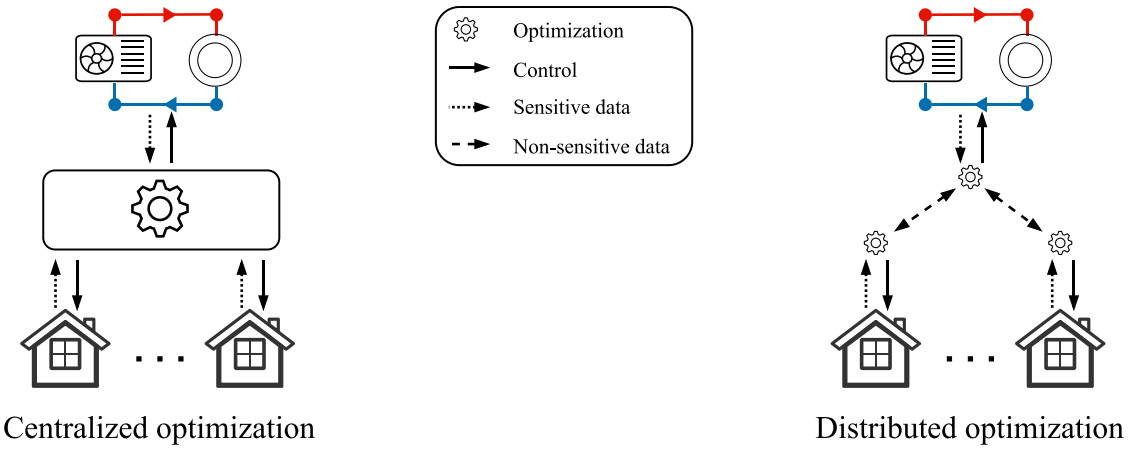


Fig. 1. Differences between centralized and distributed optimization configurations.

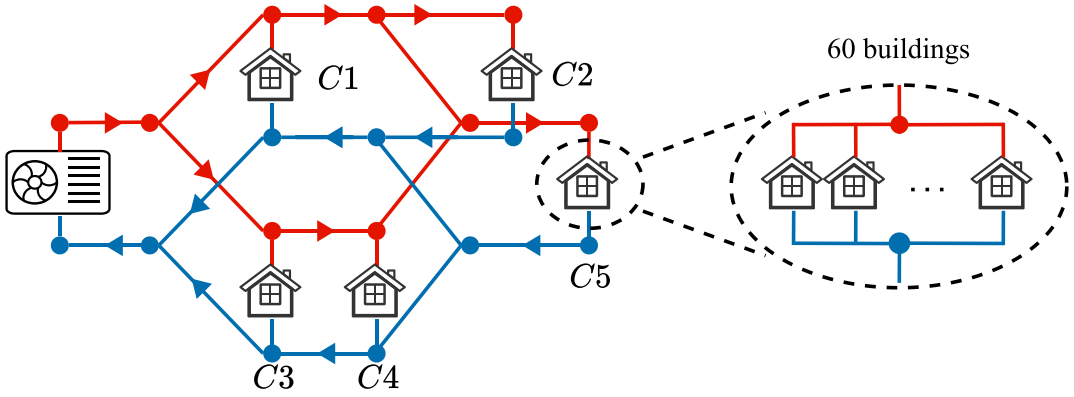


Fig. 2. The considered DHN called "AROMA" with five large-scale consumers $C1, C2, C3, C4, C5$, representing building clusters with 60 buildings each.

as a cluster of 60 individual buildings that are connected to the same nodes of the DHN. As a result, all buildings inside a cluster receive the same supply temperatures. More information on the building modeling is provided in Section 2.2.2. The consumers $C1, \dots, C5$ are referred to as "clusters" in the remainder of this work. For more detailed information about the network, the reader is referred to [15].

2.2. Control-oriented modeling

The proposed control algorithm is based on MPC and consequently requires models to predict the behavior of the considered system. This includes both a hydraulic and thermal model of the DHN and a thermal model of the buildings connected to the DHN.

2.2.1. District heating network model

The DHN is modeled as a graph consisting of nodes and edges $\mathcal{G} = (\mathcal{N}, \mathcal{E})$. Edges e include valves ($\mathcal{E}_\mathrm{v}$), pumps ($\mathcal{E}_\mathrm{pu}$), pipes ($\mathcal{E}_\mathrm{p}$), producers ($\mathcal{E}_\mathrm{pr}$), and clusters ($\mathcal{E}_\mathrm{c}$) with $\mathcal{E} = \mathcal{E}_\mathrm{v} \cup \mathcal{E}_\mathrm{pu} \cup \mathcal{E}_\mathrm{p} \cup \mathcal{E}_\mathrm{pr} \cup \mathcal{E}_\mathrm{c}$. Nodes n connect two or more edges. For any given node $n \in \mathcal{N}$, let $\mathcal{E}_n^+$ denote the edges with mass flows flowing into n and $\mathcal{E}_n^-$ the edges with mass flows exiting n. The pumps are modeled as ideal pressure sources, and the valves are designed such that all clusters receive similar mass flows.

For all edges $e \in \mathcal{E}$, conservation of mass (Eq. (1)), momentum (Eq. (2)), and energy (Eq. (3)) holds. For an incompressible, Newtonian fluid with constant density and heat capacity, these conservation equations read [22]

$$\nabla \cdot \boldsymbol{u}_e = 0 \tag{1}$$

$$\rho \left(\frac{\partial \boldsymbol{u}_e}{\partial t} + (\boldsymbol{u}_e \cdot \nabla)\, \boldsymbol{u}_e \right) = -\nabla p + \mu \nabla^2 \boldsymbol{u}_e + \boldsymbol{f} \tag{2}$$

$$\rho c_\mathrm{p} \left(\frac{\partial T_e}{\partial t} + (\boldsymbol{u}_e \cdot \nabla)\, T_e \right) = k \nabla^2 T_e + Q \tag{3}$$

with velocity vector $\boldsymbol{u}$, density ρ, dynamic viscosity μ, and heat capacity of heat carrier c_p, pressure p, heat carrier temperature T, body forces $\boldsymbol{f}$, and source term Q. The operators ∇, $\nabla\cdot$, and ∇^2 denote nabla, divergence, and Laplacian, respectively.

Pipe model

The conservation equations can be simplified under the typical assumptions of negligible friction heat and axial heat diffusion, horizontally buried pipes with a circular cross-section, and homogeneous temperatures and flows across the pipes' cross-sections [23]. Further, as pressure changes spread with the speed of sound in the DHN, while the heat carrier travels substantially slower, the hydraulics of the DHN can be modeled as stationary [24]. Based on these assumptions for pipes $e \in \mathcal{E}_\mathrm{p}$, Eqs. (2) and (3) become

$$\frac{\partial p_e}{\partial x} + f_{\mathrm{D},e} \frac{\rho}{2 d_e} |v_e| v_e = 0 \tag{4}$$

$$\frac{\partial T_e}{\partial t} + v_e \frac{\partial T_e}{\partial x} + \frac{4 U_e}{\rho c_\mathrm{p} d_e} \left(T_e - T_\mathrm{amb} \right) = 0 \tag{5}$$

where v is the heat carrier's velocity, f_D is the Darcy friction factor, which is approximated with Nikuradse's law for turbulent flow [15], d is the pipe diameter, U the heat transfer coefficient, and T_amb the ambient air temperature.

Table 1
Comparison of relevant papers from the literature proposing optimization-based controls for DHNs.

	Building flexibility	Network flexibility	Distributed optimization
Cai et al. [7]	✓	✗	✓
Qin et al. [9]	✓	✗	✓
De Giuli et al. [16]	✗	✓	✗
Hering et al. [19]	✗	✓	✗
Taylor et al. [10]	✗	✗	✓
Saletti et al. [20]	✓	✗	✗
Quagiotto et al. [21]	✗	✓	✗
Jansen et al. [11]	✓	✓	✗
Maurer et al. [8]	✗	✓	✓
Sibeijn et al. [17]	✗	✓	✗
Krug et al. [15]	✗	✓	✗
Li et al. [12]	✗	✓	✗
Present work	✓	✓	✓

Producer and cluster model

For each producer or cluster $e \in \mathcal{E}_{\text{pr}} \cup \mathcal{E}_{\text{c}}$, the heat transfer to or from the heat carrier is described as

$$c_{\text{p}} \dot{m}_e (T_{\text{in}} - T_{\text{out}}) = \phi_e \tag{6}$$

where $\dot{m} = \rho A v$ represents the mass flow of the heat carrier, where A is the cross-sectional area. Further, ϕ denotes the consumed or produced heat, and T_{in} and T_{out} are the temperatures of the inflowing and outflowing heat carrier, respectively. Pressure drops along clusters and producers are neglected.

Loops

The pressure drops in any loop $\mathcal{L}$ of the DHN have to sum up to zero as

$$\sum_{e \in \mathcal{L}} \Delta p_e = 0 \tag{7}$$

which is equivalent to Kirchhoff's voltage law in electrical networks.

Nodal model

The mass of the heat carrier is conserved by enforcing that the entering mass flows equal the exiting mass flows for each node $n \in \mathcal{N}$ in the DHN.

$$\sum_{e \in \mathcal{E}_n^+} \dot{m}_e = \sum_{e \in \mathcal{E}_n^-} \dot{m}_e \tag{8}$$

For energy conservation, the nodal temperatures T_n are calculated as flow-weighted average of the temperatures of the entering edges.

$$T_n = \frac{\sum_{e \in \mathcal{E}_n^+} \dot{m}_e \cdot T_e}{\sum_{e \in \mathcal{E}_n^+} \dot{m}_e} \tag{9}$$

2.2.2. Building models

A one-zone third-order resistance–capacitance model describes the temperature dynamics of each individual building [25]

$$C_{\text{air}} \frac{\mathrm{d}T_{\text{air}}}{\mathrm{d}t} = \frac{T_{\text{w,in}} - T_{\text{air}}}{R_{\text{w,air}}} + \frac{T_{\text{amb}} - T_{\text{air}}}{R_{\text{air,amb}}} + f_{\text{sol}} \phi_{\text{sol}} + f_{\text{conv}} \phi \tag{10}$$

$$C_{\text{w}} \frac{\mathrm{d}T_{\text{w,in}}}{\mathrm{d}t} = \frac{T_{\text{air}} - T_{\text{w,in}}}{R_{\text{w,air}}} + \frac{T_{\text{w,out}} - T_{\text{w,in}}}{R_{\text{w}}} + (1 - f_{\text{conv}}) \phi \tag{11}$$

$$C_{\text{w}} \frac{\mathrm{d}T_{\text{w,out}}}{\mathrm{d}t} = \frac{T_{\text{w,in}} - T_{\text{w,out}}}{R_{\text{w}}} + \frac{T_{\text{amb,eq}} - T_{\text{w,out}}}{R_{\text{w,amb}}}. \tag{12}$$

where f_{conv} is the convective fraction of the heat flow supplied by the DHN ϕ, f_{sol} is an identifiable parameter capturing the effect of the solar radiation ϕ_{sol} on the air temperature, C_{air} and C_{w} are the heat capacities of the air and building envelope, R the thermal resistance, and $T_{\text{amb,eq}}$ denotes the equivalent ambient temperature after accounting for the influence of solar radiation on the building envelope. It is calculated as [26]

$$T_{\text{amb,eq}} = T_{\text{amb}} + \phi_{\text{sol}} \frac{h_{\text{f}}}{h_{\text{A}}} \tag{13}$$

where $h_{\text{f}} = 0.5$ is the short-wave absorption coefficient of the exterior surface, and $h_{\text{A}} = 25\,\text{W}\,\text{m}^{-2}\,\text{K}^{-1}$ the exterior heat transfer coefficient [26].

2.3. Model predictive control

Due to the bilinear nature of heat transfer, it can be controlled by adjusting the temperature or the mass flow of the heat carrier. This leads to three possible control methods:

- Variable Flow Variable Temperature (VF-VT): both the mass flow of the heat carrier and its temperature are controlled.
- Constant Flow Variable Temperature (CF-VT): the mass flow is maintained constant and the temperature is controlled.
- Variable Flow Constant Temperature (VF-CT): the mass flow is controlled and the temperature is maintained constant.

While VF-VT allows the most granular control, it usually results in a mixed integer non-linear program when formulated as an optimal control problem. Consequently, the resulting optimization problem is typically intractable, and iterative algorithms (e.g., [12]) have to be employed to solve the optimization. However, these iterative algorithms do not guarantee convergence and can become computationally intractable even for small systems [27]. Consequently, most authors resort to simplifications such as assuming steady-state [28] or heuristically precalculating the mixed integer variables [8]. More commonly, either mass flows or temperatures are fixed, resulting in CF-VT or VF-CT formulations. Applying CF-VT or VF-CT is substantiated by their widespread application as standard operational modes in practice [24, 29]. Utilizing the heat flexibility of the heat carrier requires varying its temperature, which prohibits a VF-CT formulation. Consequently, a CF-VT formulation is adopted in the present work to leverage the heat carrier's flexibility while ensuring that the optimization problem remains tractable. A summary of challenges for optimization with VF-VT and an investigation of suitable model simplifications are provided in Appendix A.

2.3.1. Notation

Two sets are introduced for notational brevity: $\mathcal{J} = \{1, 2, \dots, J\}$ is a set containing all J flexible buildings and $\mathcal{K} = \{1, 2, \dots, N\}$ contains all natural numbers from 1 to N, where N is the number of time steps in the prediction horizon. As the MPCs use a sample time of $\Delta t = 15$ min, N equals 48 time steps for a 12-hour prediction horizon. The indices j and k denote the building and time index, respectively.

2.3.2. Model discretization

Discrete equations are required for the application in an optimization problem. Approximating $\frac{\partial p}{\partial x}$ in Eq. (4) with $\frac{\Delta p}{L_e}$ (where L_e is the pipe's length) and discretizing Eq. (5) with a first-order implicit upwind scheme, a sample time Δt and a spatial step size of Δx results in

$$\Delta p_e = L_e f_{\text{D},e} \frac{\rho}{2 d_e} |v_e| v_e \tag{14}$$

$$T_{e,k}^{w} = T_{e,k-1}^{w} + \frac{T_{e,k}^{w-1} v_e \frac{\Delta t}{\Delta x} + T_{\text{amb},k} \frac{4U_e \Delta t}{\rho c_p d_e}}{1 + v_e \frac{\Delta t}{\Delta x} + \frac{4U_e \Delta t}{\rho c_p d_e}} \tag{15}$$

where k denotes the discrete time step and w the w-th spatial element in pipe $e \in \mathcal{E}_p$. Since a CF-VT formulation is applied, neither the pressure drop Δp nor the heat carrier's velocity v changes over time.

Similarly, the building models of Eqs. (10)–(13) are discretized with a zero-order hold discretization and reformulated into a state-space system for each building j as

$$\mathbf{x}_{j,k+1} = \mathbf{A}_j \mathbf{x}_k + \mathbf{B}_j \mathbf{u}_{j,k} \tag{16}$$

$$y_{j,k} = \mathbf{C}_j \mathbf{x}_{j,k} \tag{17}$$

where the states vector contains the building's temperatures $\mathbf{x} = \begin{bmatrix} T_{\text{air}} & T_{\text{w,in}} & T_{\text{w,out}} \end{bmatrix}^{\text{T}}$, the input vector comprises the controllable heating input and the weather condition $\mathbf{u} = \begin{bmatrix} \phi & T_{\text{amb}} & \phi_{\text{sol}} \end{bmatrix}^{\text{T}}$, and the output is the indoor air temperature $y = T_{\text{air}}$.

2.3.3. Constraints

Two sets of constraints are considered: building and network constraints.

Building constraints

The MPC has to consider two constraints for each building $j \in \mathcal{J}$ for all time steps $k \in \mathcal{K}$: First, it has to maintain a comfortable indoor air temperature as indicated by time-varying and building-specific lower ($\underline{T}$) and upper ($\overline{T}$) temperature bounds.

$$\underline{T}_{j,k} \leq T_{\text{air},j,k} \leq \overline{T}_{j,k} \tag{18}$$

Second, it must consider that the heating input is limited.

$$0 \leq \phi_{\text{h},j,k} \leq \overline{\phi}_{\text{h},j} \tag{19}$$

Network constraints

Similarly, the heat pump's power consumption P_{hp} is constrained by its maximum power rating $\overline{P}_{\text{hp}}$ for $k \in \mathcal{K}$.

$$0 \leq P_{\text{hp},k} \leq \overline{P}_{\text{hp}} \tag{20}$$

The temperatures in the DHN are constrained for each time step $k \in \mathcal{K}$ with

$$T_{\text{s},k} \leq \overline{T}_{\text{s}} \tag{21}$$

$$T_{\text{r},k} \geq \underline{T}_{\text{r}} \tag{22}$$

$$T_{e,k} \geq \underline{T}_{\text{s},e} \tag{23}$$

where the operational temperature limits are defined for the supply temperature T_{s} in Eq. (21), the return temperature T_{r} in Eq. (22), and the supply temperature at each cluster T_e ($e \in \mathcal{E}_c$) in Eq. (23).

Further, the supply temperature's rate of change is constrained to avoid too high thermal stresses of the network structures [30] for $k \in \mathcal{K}$.

$$-\overline{\Delta T}_{\text{s}} \leq T_{\text{s},k+1} - T_{\text{s},k} \leq \overline{\Delta T}_{\text{s}} \tag{24}$$

Fig. 3 summarizes the control variables and constraints that are used to formulate the centralized and distributed MPCs in Section 2.3.4 and Section 2.3.5, respectively.

2.3.4. Centralized model predictive control

In the CMPC, the DHO gathers all information of the buildings and the DHN to formulate and solve a single optimization problem as

$$\min \sum_{k \in \mathcal{K}} p_{\text{el},k} \cdot P_{\text{hp},k} \tag{25}$$

subject to Eqs. (6)–(9), (14)–(24)

where p_{el} is the electricity price and P_{hp} the power consumption of the heat pump. The objective function in Eq. (25) minimizes the operating costs of the DHO while generating sufficient heat to meet the thermal demands of the clusters and the DHN's temperature constraints. For the solution of this optimization problem, the DHO must have complete knowledge about the individual buildings' system dynamics, their current temperature measurements, and their preferred occupancy (as indicated by the upper and lower temperature constraints of the buildings in Eq. (18)).

2.3.5. Distributed model predictive control

For notational brevity, the following vectors and matrices are introduced

$$\boldsymbol{\phi}_j = \begin{bmatrix} \phi_{j,1} & \cdots & \phi_{j,N} \end{bmatrix} \tag{26}$$

$$\boldsymbol{\Phi} = \begin{bmatrix} \boldsymbol{\phi}_1 \\ \vdots \\ \boldsymbol{\phi}_J \end{bmatrix} = \begin{bmatrix} \phi_{1,1} & \cdots & \phi_{1,N} \\ \vdots & \ddots & \vdots \\ \phi_{J,1} & \cdots & \phi_{J,N} \end{bmatrix} \tag{27}$$

where the vector $\boldsymbol{\phi}_j \in \mathbb{R}^{1 \times N}$ denotes the heat consumption profile along the prediction horizon of building j and the matrix $\boldsymbol{\Phi} \in \mathbb{R}^{J \times N}$ gathers the heat consumption profiles of all J buildings.

The DMPC solves the same optimization problem as the CMPC (Eq. (25)), but without the DHO receiving information about the buildings' temperature measurements, dynamics, or preferred occupancy. This is achieved by decomposing the optimization problem of Eq. (25) into $J+1$ subproblems: one for each building and one for the DHO. For this decomposition, local copies of the buildings' heat consumption $\boldsymbol{\phi}_j$ are introduced, denoted as $\widetilde{\boldsymbol{\phi}}_j$, along with the coupling constraint for $j \in \mathcal{J}$

$$\boldsymbol{\phi}_j - \widetilde{\boldsymbol{\phi}}_j = 0 \tag{28}$$

The local copies $\widetilde{\boldsymbol{\phi}}_j$ can be interpreted as the "opinion" of building j about its heat consumption, while the variable $\boldsymbol{\phi}_j$ can be considered as the DHO's "opinion" of the building's heat consumption. By applying ADMM algorithm [31], the respective "opinions" converge until Eq. (28) are satisfied. This process is illustrated in Fig. 4.

Based on the time-varying electricity price p_{el}, the DHN's temperature measurements and model, and the forecast ambient air temperature T_{amb}, the DHO computes heat consumption profiles $\boldsymbol{\phi}_j$ of the flexible buildings that minimize the heat pump's operating costs. These profiles are broadcast to the buildings. Each building then utilizes their current temperature measurements T_{air}, permitted temperature range, thermal model, and forecast ambient air temperature T_{amb} and solar radiation ϕ_{sol} to compute a heat consumption profile $\widetilde{\boldsymbol{\phi}}_j$ that follows the profile proposed by the DHO as closely as possible while satisfying the occupants' thermal comfort. The updated heat consumption profiles are relayed back to the DHO. This process is repeated iteratively until convergence is reached. While the communication requires broadcasting of heat consumption profiles, no information about occupancy, thermal dynamics, or current measurements is shared.

Mathematically, the ADMM algorithm [31] is implemented as follows. In each iteration i, the DHO solves Eq. (30), each building solves Eq. (29), and the scaled dual variable λ is updated according to Eq. (31) to drive convergence.

$$\widetilde{\boldsymbol{\phi}}_j^{i+1} = \underset{\widetilde{\phi}}{\text{argmin}} \left(\sum_{k \in \mathcal{K}} \frac{\gamma_j}{2} \| \phi_{j,k}^{i+1} - \widetilde{\phi}_{j,k} + \lambda_{j,k}^{i} \|_2^2 \right) \qquad \forall j \in \mathcal{J} \tag{29}$$

subject to Eqs. (16)–(19)

$$\boldsymbol{\Phi}^{i+1} = \underset{\phi}{\text{argmin}} \left(\sum_{k \in \mathcal{K}} p_{\text{el},k} \cdot P_{\text{hp},k} + \sum_{j \in \mathcal{J}} \frac{\gamma_j}{2} \| \phi_{j,k} - \widetilde{\phi}_{j,k}^{i} + \lambda_{j,k}^{i} \|_2^2 \right) \tag{30}$$

subject to Eqs. (6)–(9), (14), (15), (20)–(24)

$$\lambda_j^{i+1} = \boldsymbol{\phi}_j^{i+1} - \widetilde{\boldsymbol{\phi}}_j^{i+1} + \lambda_j^{i} \qquad \forall j \in \mathcal{J} \tag{31}$$

Control variables
- Heat pump power consumption P_{HP}
- Building heat consumption ϕ_j for $j \in \mathcal{J}$ buildings

Network constraints
Bounds for
- Supply temperature T_{s}
- Return temperature T_{r}
- Building cluster supply temperature T_e for $e \in \mathcal{E}_c$ clusters
- Rate of change of supply temperature ΔT_{s}
- Limited heat pump power consumption P_{HP}

Building constraints
For each building $j \in \mathcal{J}$
- Comfortable band for air temperature $T_{\mathrm{air},j}$
- Limited heat consumption ϕ_j

Fig. 3. Summary of the control variables and constraints.

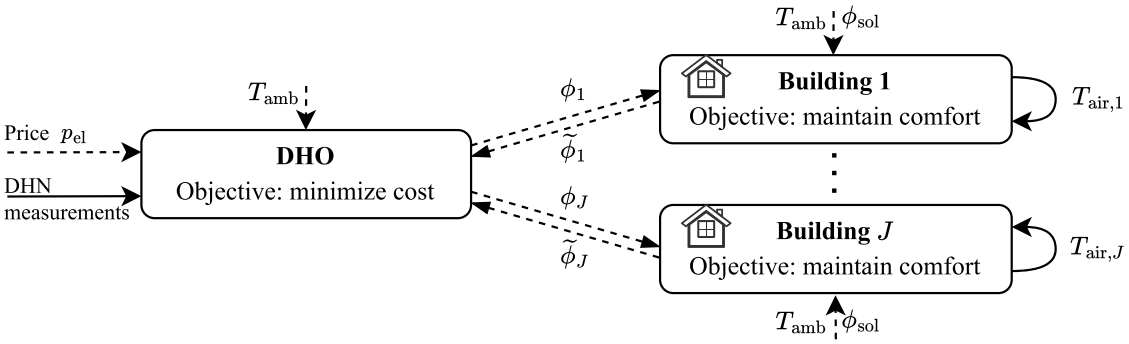


Fig. 4. The working principle of the DMPC. Solid lines represent current measurements, and dashed lines visualize information flows.

The iterations conclude when the primal (r) and dual (s) residuals fall below the tolerance ε [31]

$$\|\boldsymbol{r}_j^{i+1}\|_2 = \|\boldsymbol{\phi}_j^{i+1} - \widetilde{\boldsymbol{\phi}}_j^{i+1}\|_2 \qquad \forall j \in \mathcal{J} \tag{32}$$

$$\|\boldsymbol{s}_j^{i+1}\|_2 = \|\gamma_j \cdot \left(\widetilde{\boldsymbol{\phi}}_j^{i+1} - \widetilde{\boldsymbol{\phi}}_j^{i}\right)\|_2 \qquad \forall j \in \mathcal{J} \tag{33}$$

where the primal residual indicates the degree of satisfaction of the coupling constraints in Eq. (28) and the dual residual indicates how much the solution varies between iterations. The process of the applied ADMM algorithm is detailed in Algorithm 1.

Algorithm 1 ADMM for the distributed solution of Eq. (25)

1: Initialize $\boldsymbol{\phi}_j^0$, $\widetilde{\boldsymbol{\phi}}_j^0$, λ_j^0 $\forall j \in \mathcal{J}$
2: $i \leftarrow 0$
3: **while** $\exists j \in \mathcal{J}$ such that $\|\boldsymbol{r}_j^{i+1}\|_2 > \varepsilon_j$ **or** $\|\boldsymbol{s}_j^{i+1}\|_2 > \varepsilon_j$
4: Buildings solve Eq. (29) to obtain $\widetilde{\boldsymbol{\phi}}_j^{i+1}$ $\forall j \in \mathcal{J}$
5: DHO solves Eq. (30) to obtain $\boldsymbol{\Phi}^{i+1}$
6: Dual variables update via Eq. (31) to obtain λ_j^{i+1} $\forall j \in \mathcal{J}$
7: Update residuals $\boldsymbol{r}_j^{i+1}$ and $\boldsymbol{s}_j^{i+1}$ using Eq. (32) and Eq. Eq. (33) $\forall j \in \mathcal{J}$
8: $i \leftarrow i + 1$
9: **end**

Remark 1 (*Communication Volume*)**.** Algorithm 1 requires the exchange of the vectors $\boldsymbol{\phi}_j \in \mathbb{R}^{1\times N}$ and $\widetilde{\boldsymbol{\phi}}_j \in \mathbb{R}^{1\times N}$ (see Fig. 4) for J buildings in each iteration. Consequently, the total communication volume for a single time step is $2 \cdot N \cdot J \cdot i$ floating point numbers, where N is the prediction horizon, J is the number of flexible buildings, and i is the required number of iterations until convergence. The standard IEEE 754 [32] defines that 4 bytes is the size of a variable in the single-precision floating point format. Consequently, the total communication size of the raw exchanged data is $8 \cdot N \cdot J \cdot i$ bytes per time step. One iteration of Algorithm 1 corresponds to one communication round, in which the buildings and DHO exchange their solutions once.

Remark 2 (*Synchronization*)**.** Algorithm 1 assumes synchronization, meaning that the DHO cannot solve its local optimization (Eq. (30)) before all buildings communicate their solutions. Therefore, the algorithm's speed can be limited by the slowest communicating building.

2.4. Case study

2.4.1. Input data

A thorough evaluation of the proposed algorithm is necessary to ensure its effectiveness in various scenarios. The primary drivers in determining the optimization results are weather conditions and electricity prices. For a comprehensive evaluation, three sets of weather conditions are extracted from the Typical Meteorological Year for Karlsruhe, Germany provided by PVGIS [33]. These sets correspond to the weeks with the lowest, highest, and median ambient air temperatures during the heating season (Oct. to March). Similarly, the weeks with the highest, lowest, and median price spreads (i.e., the difference between the low and high electricity prices) during the heating season are extracted from the wholesale electricity price profile of Germany in

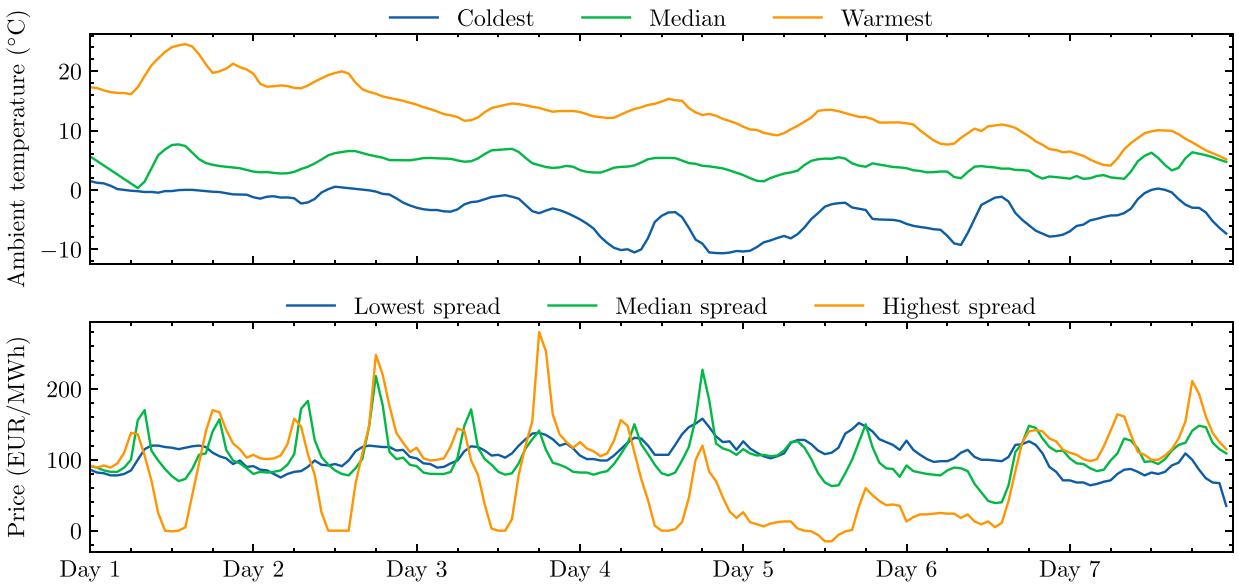


Fig. 5. The weather and price signals utilized for the evaluation.

2024, as provided by [34]. The weather and price signals are visualized in Fig. 5.

Throughout the evaluation, the weather and price forecasts are assumed to be perfect, i.e., the weather and price signals are known with certainty for the entire prediction horizon. This assumption is justified, as Section 3.5 shows that forecast errors affect DMPC and CMPC equally.

2.4.2. Buildings and loads

Two of the five clusters of the DHN (see Fig. 2) are considered as flexible and are represented by 60 building models each, leading to $J = 120$ flexible buildings. Selecting two of five clusters as flexible allows for a compromise between demonstrating the DMPC with multiple flexible clusters while maintaining computational demands manageable for a comprehensive parametric investigation. The building models are based on well-insulated and poorly insulated building models identified through measurement data from the LLEC at KIT [35,36] in [25]. From the two base models provided in [25], building models are sampled by randomly modifying each model parameter by ±25%, as also done by [37]. This procedure ensures a diverse set of residential building models with a wide range of thermal dynamics.

Based on a large-scale investigation of occupancy in residential buildings, Ueno et al. [38] propose several occupancy profiles along with their relative frequencies (i.e., how common they are). Four of these profiles are utilized to define the permitted minimum and maximum indoor air temperature profiles. These are randomly distributed across the buildings while maintaining the relative frequency of the profiles.

The thermal demand profiles of the inflexible clusters are obtained by creating a set of 60 building models and occupancy profiles, then minimizing the buildings' energy consumption for each weather scenario (i.e., cold, median, and warm).

2.5. Algorithm settings and parameters

The optimization problems are formulated using CasADi [39] and solved with Gurobi 12 [40]. Non-default Gurobi settings include: `FeasibilityTol`: 10^{-2}, `BarHomogeneous`: 1, `NumericFocus`: 2. Occasionally, the solver reports suboptimal solutions. In such cases, the optimization is restarted with modified `Presolve` settings, beginning with `Presolve`: 0 and incrementally increasing it to 1 and 2 if suboptimality persists.

Table 2
Settings for MPC and ADMM.

Variable	Value	Variable	Value
Max. iterations $i_{\max}$	100	Penalty γ	0.1
Tolerance ϵ_j	$0.01 \cdot \|\|\boldsymbol{\phi}_j\|\|_2$	Sample time Δt	15 min
Prediction horizon	12 h		

Table 2 summarizes the MPC and ADMM settings.

The maximum number of ADMM iterations $i_{\max}$ is set to 100, which is rarely reached due to a relatively loose termination tolerance ϵ_j. In particular, ϵ_j, the tolerance for the primal and dual residuals for building j, is set to $\epsilon_j = 0.01 \cdot \|\boldsymbol{\phi}_j\|_2$, which corresponds to allowing a 1% deviation between the DHO and building-level solutions. This choice ensures fast convergence without significant cost increases or comfort violations, as shown in Section 3. The penalty parameter γ is fixed at 0.1 for all buildings, unless otherwise noted.[2]

Gurobi is not warm-started with previous results, but the initial parameters $\boldsymbol{\phi}_j$, $\widetilde{\boldsymbol{\phi}}_j$, $\boldsymbol{\lambda}_j$ (see Algorithm 1) of the ADMM algorithm are initialized with the results of the previous time step shifted by one time index. For the first iteration, $\boldsymbol{\phi}_j^0$ and $\widetilde{\boldsymbol{\phi}}_j^0$ are initialized with the minimum heat consumption required to maintain thermal comfort in the buildings, while all dual variables $\boldsymbol{\lambda}_j^0$ are set to zero.

The MPCs use a prediction horizon of 12 hours and a sample time of 15 min to balance model accuracy and computation time.

The physical parameters characterizing the AROMA DHN are selected as proposed by Krug et al. [15]. While several studies use the AROMA network as a benchmark, no consistent settings for its temperature constraints have been employed. Nonetheless, in the present work, the settings are selected as similar as possible to those of other works [15,16] to improve comparability. The settings are given in Table 3.

At the start of each simulation, building temperatures are initialized to 20 °C, and the DHN water temperature is assumed to be uniformly 70 °C.

[2] While $\gamma = 0.1$ is employed as the baseline value, the robustness and optimized selection of γ will be comprehensively examined in Section 3.4.

Table 3
Parameter values for the DHN's optimization model and constraints.

Variable	Value	Variable	Value
$\overline{T}_s$	85 °C	$\Delta\overline{T}_s$	5 °C
$\underline{T}_{s,C1,\dots,C5}$	65 °C	$\underline{T}_r$	45 °C
Δx	25 m		

2.5.1. Metrics

Four metrics are used to evaluate the control performance of DMPC and CMPC: the cost increase of DMPC compared to CMPC Δ_{cost}, the buildings' average discomfort $\Delta_{\text{comf.}}$, the DHN's network constraint violations Δ_{DHN}, and the computation time. The discomfort has to be considered to ensure that the controllers do not achieve low costs by insufficiently heating the buildings.

$$\Delta_{\text{cost}} = \frac{\sum_{k=1}^{M} p_{\text{el},k} \cdot P_{\text{hp},k}^{\text{DMPC}} - \sum_{k=1}^{M} p_{\text{el},k} \cdot P_{\text{hp},k}^{\text{CMPC}}}{\sum_{k=1}^{M} p_{\text{el},k} \cdot P_{\text{hp},k}^{\text{CMPC}}} \tag{34}$$

$$\Delta_{\text{comf.}} = \frac{1}{J} \cdot \Delta t \cdot \sum_{j=1}^{J} \sum_{k=1}^{M} \left(\left[\underline{T}_{j,k} - T_{\text{air},j,k} \right]^+ + \left[T_{\text{air},j,k} - \overline{T}_{j,k} \right]^+ \right) \tag{35}$$

$$\Delta_{\text{DHN}} = \Delta t \cdot \sum_{k=1}^{M} \left(d_k^{\text{DMPC}} - d_k^{\text{CMPC}} \right) \tag{36}$$

$$d_k = \underbrace{\left[T_{s,k} - \overline{T}_s \right]^+}_{\text{(i)}} + \underbrace{\left[\underline{T}_r - T_{r,k} \right]^+}_{\text{(ii)}} + \underbrace{\sum_{e \in \mathcal{E}_c} \left[\underline{T}_{s,e} - T_{e,k} \right]^+}_{\text{(iii)}} + \underbrace{\left[\left| T_{s,k+1} - T_{s,k} \right| - \overline{\Delta T}_s \right]^+}_{\text{(iv)}} \tag{37}$$

Eq. (34) calculates the cost increase in the heat pump's operation under DMPC compared to CMPC across the simulation time M. The buildings' average discomfort is quantified with Eq. (35), in which discomfort is defined as violating the permitted temperature interval defined by $[\underline{T}, \overline{T}]$, where the operator $[\cdot]^+$ is a shorthand for $\max(0, \cdot)$. The increase of violations of the DHN constraints under DMPC compared to CMPC is quantified in Eq. (36). The DHN constraint violations d consist of constraint violations of (i) the supply temperature, (ii) the return temperature, (iii) the consumers' supply temperatures, and (iv) the supply temperature's permitted rate of change.

For the CMPC, the computation time is the required time to solve the centralized optimization problem. In contrast, the DMPC's computation time t_{DMPC} is calculated as

$$t_{\text{DMPC}} = t_{\text{DHO}} + \frac{1}{J} \sum_{j \in \mathcal{J}} t_j \tag{38}$$

where t_{DHO} is the time required to solve the optimization problem of the DHO, and t_j is the time required to solve the optimization problem of building j. The division by J (the number of buildings) accounts for the fact that all buildings' optimizations are solved in parallel. In a real application of the DMPC, additional time for communication between DHO and buildings is required, however, this has been shown as negligible by Cai et al. [7].

2.6. Emulator model

In real applications, forecast errors due to imperfect models are inevitable, resulting in suboptimal control behavior. To evaluate whether the presented control can successfully handle deviations from the optimal control trajectory, the control inputs are fed into an emulator model of the real DHN. This model is more complex and accurate than the simplified control-oriented model presented in Section 2.2.1.

Previous studies have successfully developed high-fidelity emulator models for the AROMA DHN using the Modelica package *DHN4Control* [16,41]. Consequently, the same procedure is applied to develop the emulator model for this work. For detailed information about the underlying models of *DHN4Control*, please refer to [41].

Table 4
Absolute and relative savings for one week due to building and network flexibility compared to "Baseline" without flexibility. "Network" uses network flexibility only, "Building and network" uses both flexibility sources.

Price spread	Weather	Network		Building and network	
		Savings	Savings	Savings	Savings
Low	Cold	0.53%	47.8 €	1.79%	161.6 €
Low	Median	1.83%	114.3 €	2.82%	176.5 €
Low	Warm	4.40%	107.5 €	6.11%	149.2 €
Median	Cold	2.46%	219.8 €	5.50%	491.4 €
Median	Median	5.67%	354.3 €	11.54%	721.5 €
Median	Warm	10.58%	267.7 €	20.55%	519.8 €
High	Cold	5.04%	350.7 €	12.76%	887.9 €
High	Median	9.72%	489.1 €	26.81%	1348.6 €
High	Warm	19.13%	398.0 €	51.28%	1066.6 €

3. Results and discussion

3.1. Economic benefit of building flexibility

Centralized optimization requires access to privacy-critical information to utilize building flexibility, which may prevent its application in real-world scenarios. In contrast, distributed optimization can leverage building flexibility without privacy-critical information. To quantify the potential for cost savings of the DMPC, the optimization results of three configurations are evaluated:

1. "Baseline": A price-agnostic optimization using neither building nor network flexibility. Building flexibility is removed by minimizing the buildings' energy consumption, while network flexibility is removed by maintaining a constant supply temperature of 75°C.
2. "Network": Network flexibility is enabled by allowing variable supply temperatures, while the buildings' flexibility remains unused. This approach utilizes the heat carrier as thermal storage to shift loads to low-price periods. This configuration is suited for centralized optimization, as it does not require privacy-critical building data.
3. "Building and network": Both building and network flexibility are leveraged for load shifting. This configuration may only be achievable through distributed optimization in real-world scenarios.

Table 4 presents the resulting cost savings for "Network" and "Building and network" compared to "Baseline" under varying price spreads and weather conditions (see Fig. 5). The cost savings increase considerably with the price spread. The relative savings are highest in warm weather and lowest in cold weather, as the high heating demand during cold weather forces continuous heat pump operation even during high-price periods. Conversely, the absolute savings peak under median weather conditions. While network flexibility alone already achieves cost savings between 0.53% and 19.13%, adding building flexibility approximately doubles the savings, reaching up to 51.28%. This shows that the optimal operation of DHNs using network flexibility is promising, but distributed optimization offers substantial additional potential by unlocking building flexibility.

3.2. Convergence of the DMPC

Since the proposed DMPC relies on iterative communication between the DHO and the participating clusters, it has to be ensured

Table 5
Impact of different flexible clusters on the DMPC's performance compared to the CMPC under median weather and median price spread.

Flexible clusters	Δ_{cost}	$\Delta_{comf.}$	Δ_{DHN}	Comp. time
$(C1, C2)$	0.62% (+35 €)	0.04 Kh	1.37 Kh	−67.8%
$(C1, C3)$	0.69% (+39 €)	0.04 Kh	−6.28 Kh	−74.6%
$(C1, C4)$	0.56% (+32 €)	0.04 Kh	−1.05 Kh	−64.5%
$(C1, C5)$	0.71% (+40 €)	0.04 Kh	8.47 Kh	−65.7%
$(C2, C3)$	0.64% (+36 €)	0.04 Kh	−2.39 Kh	−45.9%
$(C2, C4)$	0.72% (+41 €)	0.04 Kh	9.40 Kh	−58.7%
$(C2, C5)$	0.80% (+45 €)	0.04 Kh	4.04 Kh	−59.5%
$(C3, C4)$	0.60% (+34 €)	0.04 Kh	−0.41 Kh	−59.2%
$(C3, C5)$	0.66% (+38 €)	0.04 Kh	0.67 Kh	−69.7%
$(C4, C5)$	0.71% (+40 €)	0.04 Kh	3.86 Kh	−70.0%
$(C1, C2, C3, C4, C5)$	1.11% (+61 €)	0.04 Kh	0.47 Kh	−80.4%

that no convergence issues arise across a wide range of possible scenarios. This concern is addressed in two experiments. First, all possible combinations of drawing two flexible clusters from the five clusters are evaluated under the most representative conditions, i.e., using median price spreads and median weather data. For completeness, the case when all five clusters are flexible is also evaluated. Table 5 presents the results of this simulation.

The DMPC consistently achieves costs and thermal comfort comparable to that of the CMPC benchmark. Specifically, the excess costs incurred by the DMPC range from 0.56% to 1.11%, with negligible additional discomfort of 0.04 Kh per building. Regarding the network constraint violations Δ_{DHN}, the DMPC performs similarly to the CMPC, neither systematically increasing nor decreasing the violations. Further, the DMPC consistently achieves computation time savings between 45.9% and 80.4%. The DMPC performs similarly with five flexible clusters as it does with two flexible clusters. While the cost increase is marginally higher, the computation time savings increase as well, since the computational load is distributed across more participants. In summary, this experiment results in two findings: (i) the DMPC achieves a similar control performance to the CMPC while reducing computation time and respecting privacy constraints, and (ii) the DMPC is insensitive to the selection of participating clusters. Consequently, an arbitrary pair of flexible clusters $(C1, C5)$ is selected for more detailed analysis.

In the second experiment, an extensive evaluation of the DMPC is conducted by investigating all possible combinations of price and weather data for the selected pair of flexible clusters. Table 6 presents the results.

The DMPC achieves performance similar to the CMPC across all scenarios, indicating robustness to variations in price and weather data. The discomfort is negligible in each scenario, ranging from 0.04 Kh to 0.07 Kh per building. In eight out of nine scenarios, the minor cost increase ranges from 0.09% to 0.71%. For the scenario with a high price spread and warm weather, the DMPC even slightly reduces costs compared to the CMPC benchmark by 1.11%. Although the CMPC typically obtains the best possible results, this outcome can occur for two reasons: first, slightly suboptimal results from DMPC can outperform the results of CMPC due to model inaccuracies; second, the larger centralized optimization problem is harder to solve than the smaller distributed counterparts, which can lead to suboptimal termination of the optimization solver for the CMPC.

While the DMPC consistently reduces the computation time across all combinations of price and weather data, the savings vary between 15.5% and 74.3% compared to the CMPC. Fig. 6 shows the correlation between the computation time savings and its main driver: the number of iterations required for the DMPC to conclude.

Notably, the DMPC converges quickly within three to five mean iterations across all cases, which is achieved by warm-starting the DMPC's ADMM algorithm with the results of the previous time step. Further, Fig. 6 demonstrates a clear correlation ($R^2 = 0.83$) between the average iterations and the computation time savings. Notably, two additional mean iterations reduce the computation time savings from 65% to 30%. Consequently, avoiding excess iterations is essential for the practical application of DMPC. Although the distribution of iterations across scenarios shows that the algorithm converges within five iterations for the vast majority of cases, two outliers reach the maximum iteration limit of 100 iterations.

In summary, the DMPC achieves robust performance and negligible comfort violations across all combinations of price spreads and weather conditions. While the control performance varies only slightly, it varies more than in the previous experiment (see Table 5), indicating that the price spreads and weather conditions have a more substantial impact on the performance than the location of the flexible clusters.

3.3. Demonstration of controller behavior

Fig. 7 presents the heat pump power consumption for a representative day in cold and warm weather, along with the median price spread for both the CMPC and DMPC. Notably, both controls take very similar, although not identical, control actions. The minor difference is due to the tolerance of the ADMM algorithm, which aborts the iterative process when the primal and dual residuals fall below the accepted tolerances, leading to minor excess costs of 0.46%, and 0.41%, respectively (see Table 6). In general, the heat pump's power consumption is shifted to periods with low electricity prices (e.g., 01:00 until 06:00). While electricity is more expensive (e.g., 07:00 until 10:00), the controllers avoid heating. Clearly, load shifting is successful in warm weather as heating can be almost entirely avoided during periods of higher electricity prices. While the heat pump's operation during cold weather is also primarily shifted to low-price periods, the heat demand is so high that it still has to generate heat during high-price periods, albeit at a lower capacity.

Fig. 8 presents how the CMPC and DMPC utilize the heat carrier as thermal storage. The upper plot shows the supply and return temperatures at the heat pump for CMPC and DMPC, the middle plot shows the supply temperatures at the flexible clusters, and the bottom plot shows the electricity price.

Three statements can be deduced from Fig. 8: First, the thermal storage capacity of the heat carrier is used by "charging" it (i.e., heating it) extensively during low-price times and discharging it during high-price periods. Second, the temperature constraints of the supply and return temperatures are barely violated, indicating sufficient model accuracy. Third, the solutions of the CMPC and DMPC are nearly identical, indicating that the DMPC successfully approximates the control behavior of the CMPC.

In addition to using the heat carrier as thermal storage, the controls can also utilize the flexible buildings as storage. This is visualized in Fig. 9, which presents indoor air temperatures of four exemplary buildings with different temperature constraints when using DMPC under median weather and high and low price spreads.

Notably, both the minimum and maximum temperature constraints are respected, as already indicated in Table 6 by the negligible discomfort. While the DMPC minimizes the operating cost of the DHO in both cases with high and low price spreads, respectively, the strategy differs considerably. When the price spread is low, using the buildings' thermal mass for heat storage is not cost-effective, so the cost minimization reduces to energy minimization. In contrast, the buildings' thermal flexibility is fully utilized when the price spread is high, as evident from the considerable preheating. When a CMPC cannot utilize the building's flexibility due to privacy concerns in a real application, the DMPC approach is especially beneficial in markets with high price spreads to unlock additional storage capacity.

Table 6
Impact of varying price spread and weather on the DMPC's control performance compared to the CMPC for the pair of flexible clusters $(C1, C5)$.

Price spread	Weather	Δ_{cost}	$\Delta_{\text{comf.}}$	Δ_{DHN}	Comp. time
Low	Cold	0.09% (+7 €)	0.07 Kh	−4.23 Kh	−15.5%
Low	Median	0.31% (+19 €)	0.05 Kh	9.28 Kh	−58.9%
Low	Warm	0.10% (+2 €)	0.07 Kh	−5.78 Kh	−27.2%
Median	Cold	0.46% (+34 €)	0.06 Kh	−6.57 Kh	−58.9%
Median	Median	0.71% (+40 €)	0.04 Kh	8.48 Kh	−65.7%
Median	Warm	0.41% (+9 €)	0.05 Kh	−2.95 Kh	−44.9%
High	Cold	0.24% (+12 €)	0.04 Kh	4.36 Kh	−64.5%
High	Median	0.62% (+24 €)	0.03 Kh	−9.79 Kh	−74.3%
High	Warm	−1.11% (−15 €)	0.05 Kh	−13.45 Kh	−51.7%

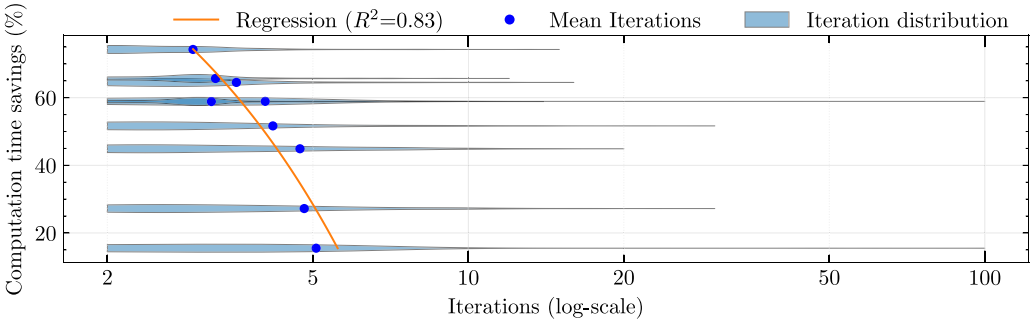
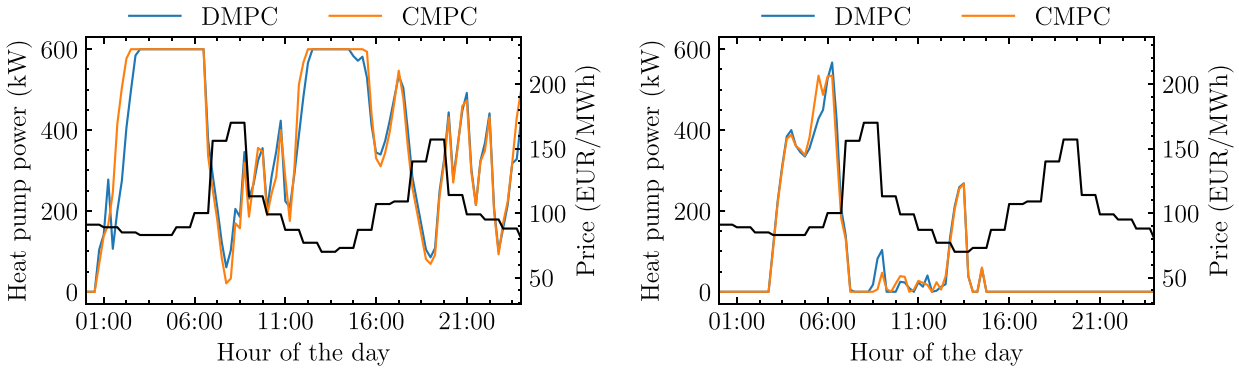


Fig. 6. Distributions of the iterations it takes DMPC to converge and correlation between the computation time savings and the mean iterations for the DMPC to conclude.



(a) Median price spread, cold weather.

(b) Median price spread, warm weather.

Fig. 7. The heat pump power consumption and electricity price for CMPC and DMPC in two different scenarios.

3.4. Scalability of DMPC

This section explores the scalability of DMPC and the impact of different ADMM parameter settings on convergence and optimality. The settings include varying values of the penalty parameter γ for both cold starting and warm starting, and different numbers of flexible buildings $J \in \{120, 1200\}$.

The performance of the ADMM algorithm is evaluated based on three criteria:

- The primal residual $\sum_j \|\mathbf{r}_j\|_2$, which indicates the violation of the coupling constraints (see Eq. (28)).
- The dual residual $\sum_j \|\mathbf{s}_j\|_2$, which reflects how much the solution varies between iterations.
- The optimality gap of the ADMM solution (x) compared to the optimal central solution $(x^\star)$, calculated as $\frac{f(x)-f(x^\star)}{f(x^\star)}$, where $f(x)$ and $f(x^\star)$ denote the values of the objective functions.

Since the optimal central solution is usually unknown, the ADMM algorithm terminates when the primal and dual residuals fall below specified tolerance thresholds (compare Algorithm 1), without considering optimality. Notably, low residuals do not automatically guarantee a solution near the optimal solution, as the algorithm may also converge to suboptimal solutions.

Fig. 10 illustrates the convergence and optimality of the ADMM algorithm for various values of γ, where γ remains constant for each building throughout the iterations, in addition to an adaptive tuning strategy for γ [31]. This strategy aims to balance the primal and dual residuals as

$$\gamma_j^{i+1} = \begin{cases} 2\gamma_j^i & \text{if } \|\mathbf{r}_j\|_2 > 10\|\mathbf{s}_j\|_2 \\ \frac{1}{2}\gamma_j^i & \text{if } \|\mathbf{s}_j\|_2 > 10\|\mathbf{r}_j\|_2 \\ \gamma_j^i & \text{otherwise} \end{cases} \tag{39}$$

where γ_j^{i+1} is the adapted penalty of iteration $i+1$ for building j, based on the proportions of the primal and dual residuals for this building

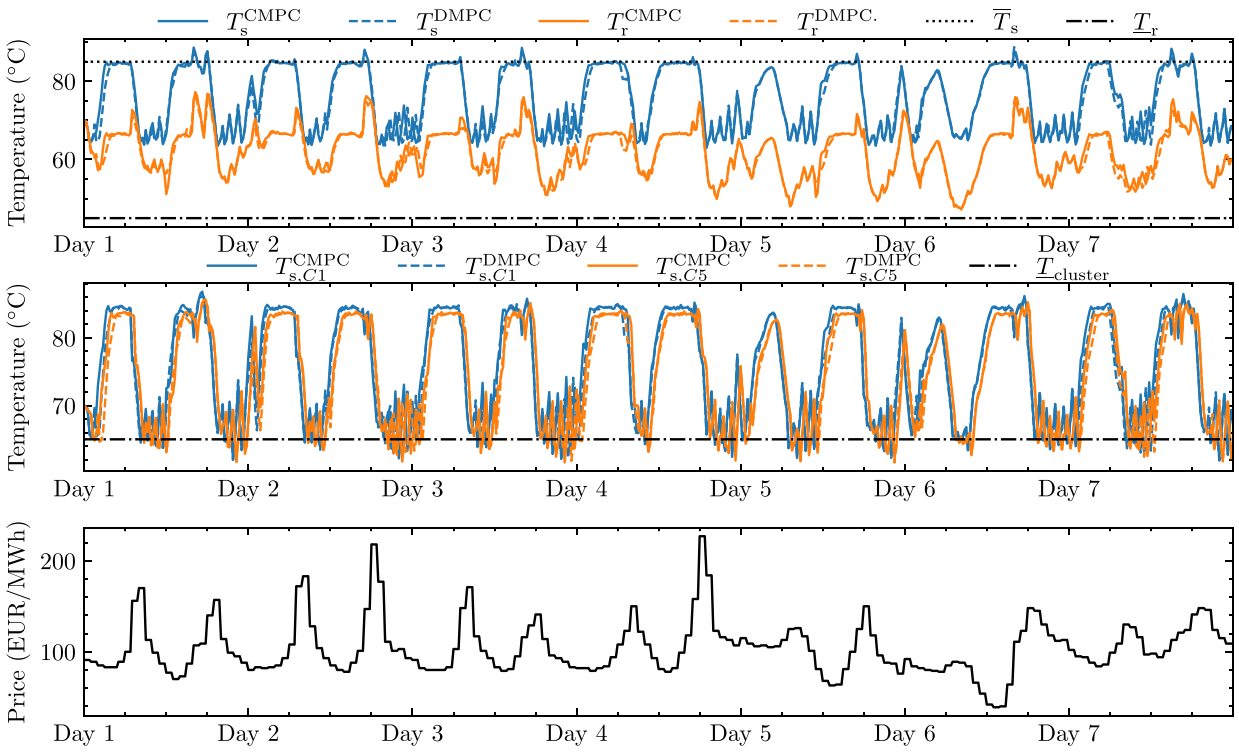


Fig. 8. Utilization of the heat carrier as heat storage for median price and cold weather.

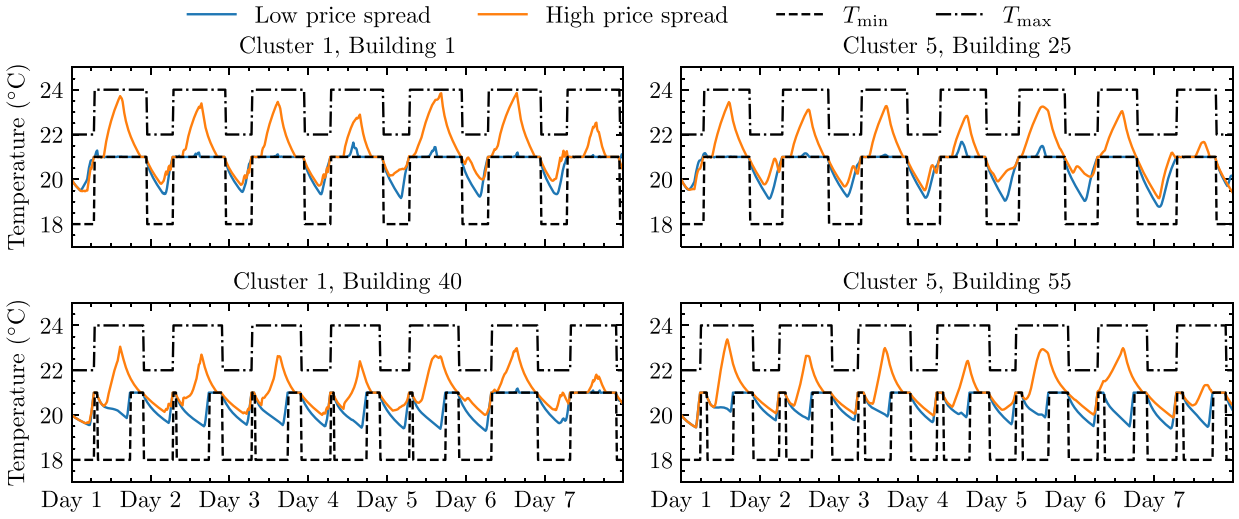


Fig. 9. Indoor air temperatures under DMPC for median weather and low and high price spreads.

in iteration i. The black dashed line marks the tolerance of the ADMM algorithm as discussed in Section 2.5.

The DHO's objective is to minimize the heat pump operating cost while the buildings aim to maintain thermal comfort. The penalty γ incentivizes each participant to consider the objectives of the other participants. If $\gamma = 0$, the DHO would only minimize operating costs, disregarding the buildings' requested heating. The buildings, in turn, could not accept the DHO's proposed solution, as it would violate their temperature constraints, resulting in a deadlock.

For small values of γ (see $\gamma = 0.01$), the incentive to consider the objectives of other participants is weak, resulting in large primal residuals (indicating significant discrepancies between the DHO's and the buildings' solutions) and low dual residuals (reflecting low incentive to adjust the solution), which hinders convergence. In contrast, large values of γ (see $\gamma = 1$), impose strong penalties for disagreement between the DHO and the buildings, leading to low primal residuals but slowly decaying dual residuals, which also inhibits convergence. This strong incentive for agreement between DHO and buildings can result in suboptimal solutions, as shown by the optimality gap in Fig. 10. In contrast, the ADMM with both the adaptive and the well-tuned penalty ($\gamma = 0.2$) exhibits balanced primal and dual residuals, allowing quick convergence and rapidly reaching optimality gaps < 1%.

Fig. 11 shows the convergence and optimality gap of the ADMM algorithm with cold and warm starts applied for $J = 120$ flexible buildings with $\gamma = 0.1$, and $J = 1200$ flexible buildings with $\gamma = 0.2$. These values of γ have been identified as optimal for the respective numbers of flexible buildings.

When cold-starting the ADMM algorithm, the initial residuals are large. For $J = 1200$ buildings, the residuals converge more slowly than they do for $J = 120$, taking ten iterations to drop below the tolerance

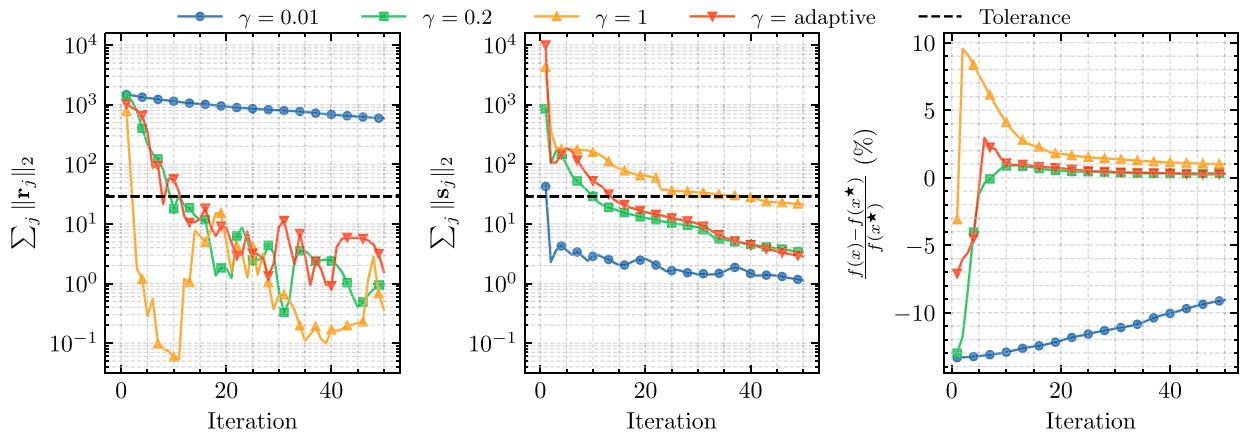


Fig. 10. Convergence of the ADMM algorithm for 1200 flexible buildings with varying penalty parameter γ as indicated by primal residual, dual residual, and optimality gap.

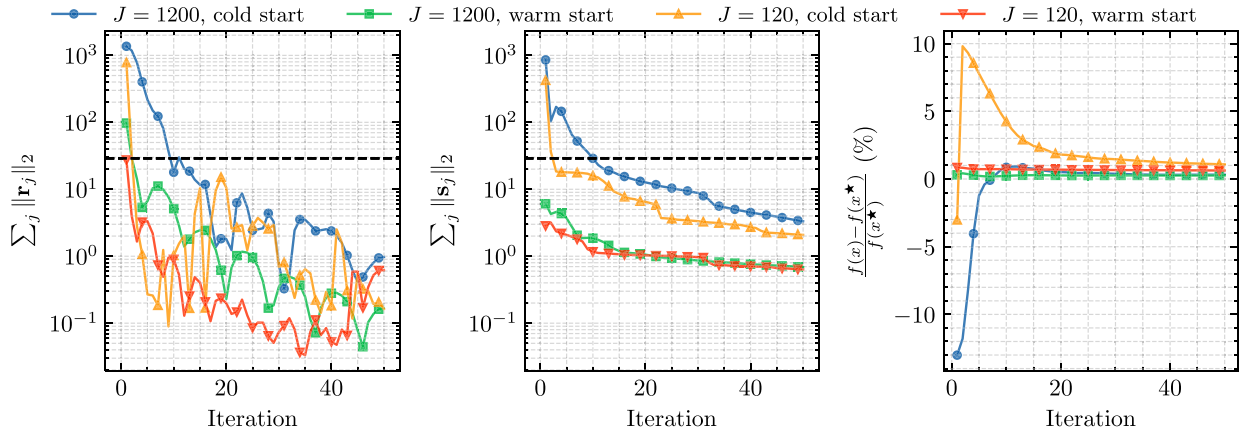


Fig. 11. Convergence of the ADMM algorithm for 120 and 1200 flexible buildings with cold and warm starts.

compared to four iterations for 120 flexible buildings. Conversely, the optimality gap improves more rapidly for the 1200 flexible buildings. In fact, the ADMM algorithm with a cold start would have terminated already with a substantial suboptimality > 5% for the case with 120 flexible buildings.

In contrast, when the algorithm is warm-started, the residuals drop below the tolerance within just two iterations while achieving optimality gaps below 1%. This highlights the importance of warm-starting the ADMM algorithm to avoid excess iteration and to achieve low optimality gaps. Overall, the algorithm seems insensitive to the increase from 120 to 1200 flexible buildings, indicating that it scales well. However, different values of γ may yield optimal results for different numbers of flexible buildings. To avoid time-consuming tuning of γ, an adaptive update (see Eq. (39)) can be applied.

Fig. 12 compares the computation times per time step for DMPC and CMPC for 120 and 1200 flexible buildings to assess the scalability of the DMPC. The computation times for DMPC are calculated according to Eq. (38). The left plot conveys two key points. First, the computation time for the DMPC is primarily determined by the time required to optimize the DHN, while the parallel optimization of the buildings is executed almost immediately. This suggests that optimizing the DHN might be the bottleneck when scaling the approach to large-scale DHNs. Second, as the number of flexible buildings increases, the computation time for CMPC increases significantly more than the computation time for DMPC, demonstrating the scalability benefits of distributed optimization. Specifically, the median computation time savings increase from 58% for 120 flexible buildings to 88% for 1200 flexible buildings.

The right plot displays the distribution of the computation times on a logarithmic scale. Both DMPC and CMPC exhibit outliers with substantial computation times of multiple minutes per time step. For DMPC, these outliers occur when a high number of iterations is required until the termination criteria are satisfied. For CMPC, outliers occur when the solver takes an excessive amount of time to find the optimal solution. In both cases, these outliers can be mitigated in real applications by setting a maximum time limit for DMPC or CMPC, at the cost of suboptimal solutions.

3.5. Robustness of DMPC under imperfect forecasts

Three sources of uncertainty are considered to evaluate the DMPC's robustness under erroneous forecasts:

1. Heat load uncertainty: Five studies [42–46] on day-ahead heat load forecasting (from building to DHN level) report MAPE or CV-RMSE values between 5% and 15%. However, none provide details on the distribution of forecast errors. Therefore, a uniform error distribution with forecast errors in the range of ± 20% is assumed.
2. Ambient air temperature uncertainty: Forecast errors are modeled as Gaussian with zero mean and a standard deviation of 1.61 K [25].
3. Solar radiation uncertainty: The forecast errors are approximated as zero at night and Gaussian during the day, with zero mean and a standard deviation of 69 W/m^2 during the day [47].

Latin Hypercube Sampling is used to generate 100 scenarios of forecast error combinations for heat load, ambient air temperature, and solar radiation. DMPC and CMPC are then simulated for one day per

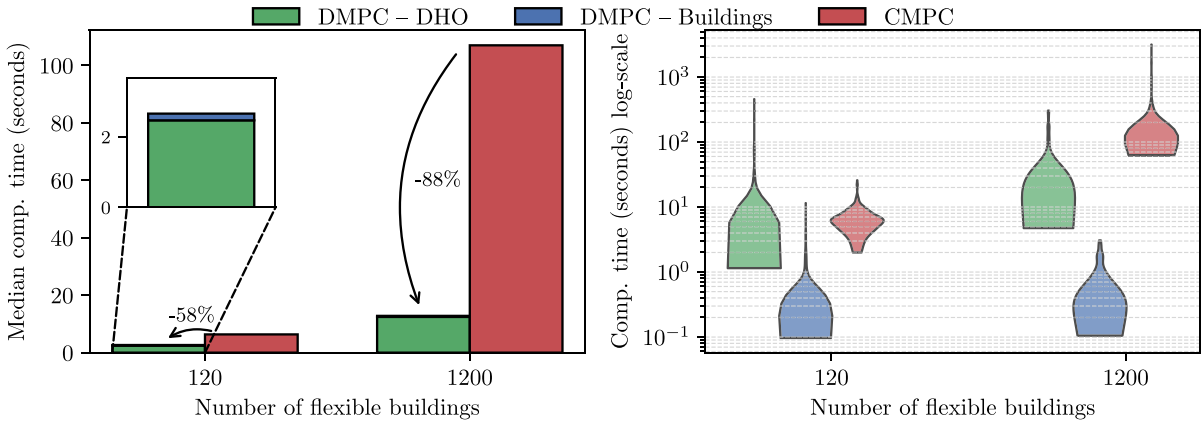


Fig. 12. Computation times per time step for the DMPC (for buildings and DHO) and CMPC for 120 and 1200 flexible buildings, respectively. The left plot shows the median computation times and the right plot the distributions on a log-scale.

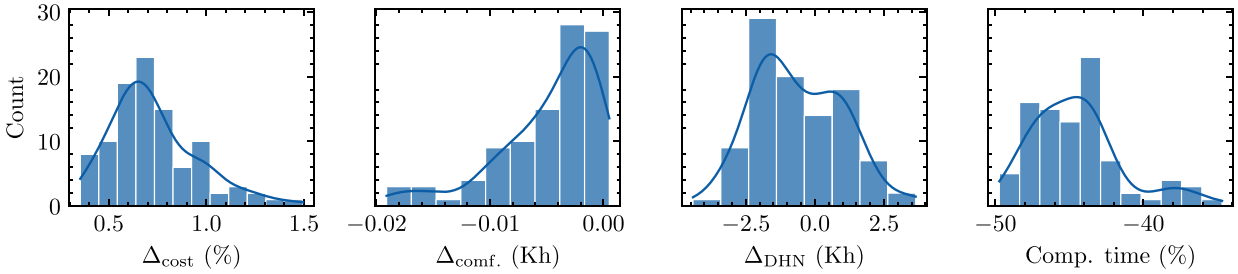


Fig. 13. Increases in cost Δ_{cost}, comfort violations $\Delta_{\text{comf.}}$, network constraint violations Δ_{DHN}, and computation time savings of DMPC compared to CMPC in 100 simulations with varying forecast errors.

scenario using median weather data as "true" weather data and the sampled forecast errors. Fig. 13 presents the distributions of the DMPC's achieved performance.

The DMPC shows robust performance under imperfect forecasts, achieving similar performance across all 100 scenarios. In particular, the average cost increase (Δ_{cost}) is 0.7%, and the maximum observed cost increase is 1.5%. Comfort violations ($\Delta_{\text{comf.}}$) remain almost identical to CMPC, and network constraint violations (Δ_{DHN}) are centered around zero, indicating no systematic improvement or deterioration of the DHN control compared to CMPC. Further, the DMPC consistently achieves computation time savings of 40% to 50% in almost all cases.

In summary, the results show that the DMPC is as robust to uncertain forecasts as the CMPC, achieving almost identical control performance while maintaining privacy and substantially reducing computation time.

4. Conclusion

The integration of large-scale heat pumps into District Heating Networks (DHNs) can decarbonize heating supply while utilizing the DHNs' substantial potential to provide demand flexibility for the power grid. This flexibility is increasingly important for balancing the power grid in light of an increasing share of renewable energy sources. To operate the heat pumps in response to grid signals, advanced controls are required, such as Model Predictive Control (MPC).

The present work introduces a Distributed Model Predictive Control (DMPC) method for the demand response of DHNs based on the Alternating Direction Method of Multipliers (ADMM) algorithm. The DMPC is designed to enable demand response without relying on an omniscient centralized controller, thus preserving privacy, fostering scalability, and reducing computational requirements. While previous studies have either neglected the flexibility potential of the heat carrier or the thermal mass of buildings, the proposed DMPC utilizes both to adjust the heat generation of a large-scale heat pump in response to dynamic pricing.

The DMPC is evaluated in nine one-week simulation scenarios, characterized by variations in price spreads and weather conditions. The results demonstrate that the DMPC achieves nearly identical performance to the centralized MPC (CMPC) benchmark in terms of cost and thermal comfort, with excess costs below 1.5% and negligible discomfort. Additionally, the DMPC consistently reduces computation time by 15.5% to 74.3% compared to CMPC, enhancing scalability and making it suitable for larger networks. The computation time savings are substantially influenced by the number of iterations required for the DMPC to converge, highlighting the importance of avoiding excess iterations. The method is robust to variations in both the selection of flexible clusters and weather conditions, and is particularly beneficial in markets with high price spreads, where the utilization of building flexibility can substantially contribute to cost savings.

The present study utilizes a synchronous ADMM algorithm to coordinate the buildings with the district heating operator. However, if one building experiences delays, it can slow down the entire algorithm. Therefore, future research may investigate the potential of an asynchronous ADMM approach. Additionally, the DHN model relies on accurate knowledge of the physical parameters of the network pipes, which may not always be available in practice. This issue can be addressed by exploring data-driven modeling techniques, which can also aid in simplifying the nonlinearity of the optimization problem. Future work will explore the potential for combining DMPC with other distributed energy resources, such as photovoltaic systems or battery storage, to maximize the overall flexibility and efficiency of future DHNs.

CRediT authorship contribution statement

Felix Langner: Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Jörg**

Matthes: Writing – review & editing, Supervision, Resources, Funding acquisition, Conceptualization. **Veit Hagenmeyer:** Writing – review & editing, Supervision, Resources, Funding acquisition.

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

During the preparation of this work the authors used Microsoft Copilot in order to improve the readability and language of the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Declaration of competing interest

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

Acknowledgments

We acknowledge the financial support of the Helmholtz Association of German Research Centres (HGF) within the framework of the Program-Oriented Funding POF IV in the program Energy Systems Design (ESD, project number 37.12.01).

Appendix A. Challenges in optimizing VF-VT

Operating DHNs with varying mass flows and temperatures offers the most granular control but results in a complex, often intractable optimization problem. This section outlines the main challenges and presents model simplifications (Appendices A.1.1 to A.1.4) that can be applied to simplify the optimization. Finally, Appendix A.2 assesses the feasibility and suitability of these model simplifications for MPC applications. The challenges in optimizing VF-VT systems include:

(C1) In meshed networks, such as the AROMA network, the flow direction in each edge is not known a priori and can switch depending on valve settings. This has significant implications for the optimization:

- The pressure drop cannot be modeled as $\Delta p \propto \dot{m}^2$ but has to be expressed as $\Delta p \propto |\dot{m}|\dot{m}$ because the sign of the pressure drop changes if the flow direction switches. This formulation for Δp is non-differentiable at $\dot{m} = 0$, introducing non-smoothness into the optimization.
- When flows can change direction, formulating the mixing equation for the nodes (see Eq. (9)) becomes complicated. This is because the formulation of the mixing equations requires knowing sets of edges with flows entering and exiting the mixing node. However, these sets result from solving the optimization and are unknown beforehand.

(C2) Even in tree networks where the flow direction for each edge is known a priori, the optimization problem remains non-convex due to a second-order equality constraint $\Delta p \propto \dot{m}^2$ and Kirchhoff's loop law (Eq. (7)).

(C3) The heat transfer equation (Eq. (6)) $\phi = c\dot{m}\Delta T$ is bilinear in $\dot{m}$ and ΔT, adding further nonlinearity to the optimization problem.

(C4) Computing the thermal transients in the edges adds complexity and nonlinearity. The widespread "node method" uses auxiliary variables to track the hot-water front.

A.1. Model simplifications

A.1.1. Unknown flow direction

Even in meshed networks, the flow direction is determined a priori for most edges, and a preprocessing algorithm (e.g., [15]) can be applied to identify them. This motivates two different simplifications for challenge (C1):

1. The flow direction can be predetermined for the optimization by assuming a certain flow direction [48]. This approach lowers computational costs at the cost of reducing the feasible set for the optimization problem.
2. Alternatively, to retain the ability to account for switching flow directions, the mass flows of edges with potentially bidirectional flows can be decomposed as $\dot{m} = \dot{m}^+ - \dot{m}^-$ where $\dot{m}^+, \dot{m}^- \geq 0$. The complementary constraint $\dot{m}^+ \cdot \dot{m}^- = 0$ ensures that only one of these flows is active at a time (see, e.g. [17]). With this reformulation, the pressure loss is calculated as $\Delta p \propto (\dot{m}^+ \cdot \dot{m}^+ - \dot{m}^- \cdot \dot{m}^-)$. However, implementing complementary constraints significantly increases the complexity of the optimization problem.

Decomposing the mass flow into its positive and negative components ($\dot{m} = \dot{m}^+ - \dot{m}^-$) also allows a natural formulation of the mixing equations by simply summing both negative and positive flows, while the complementary constraint drives one of the flows to zero. However, as complementary-constrained programs are hard to solve, Krug et al. [15] propose a reformulation based on nonlinear programming.

A.1.2. Convex reformulation of Kirchhoff's loop law

Non-convex optimization problems are highly challenging to solve because they contain local minima. Consequently, it is desirable to convexify the optimization problem to address challenge (C2). The strategy to convexify Kirchhoff's loop law, briefly summarized here for clarity, has been proposed by Sibeijn et al. [17]. More detailed information can be found in [17].

The main idea is to replace the non-convex squared *equality* constraint that models pressure losses (see Eq. (14)) by convex squared *inequality* constraints. In any loop $\mathcal{L}$, the maximum possible pressure head that the pump can generate imposes an upper limit on the pressure drop across this loop as

$$\sum_{e \in \mathcal{L}} \mu_e \dot{m}_e^2 \leq \Delta p_{\mathrm{p,max}} \tag{A.1}$$

which corresponds to a network configuration where all valves are fully open and the pump operates at its maximum capacity [17]. In Eq. (A.1), μ_e denotes the friction coefficient and $\Delta p_{\mathrm{p,max}}$ the maximum pressure increase the pump can supply.

However, simply by imposing Eq. (A.1) as a constraint in the optimization, Eq. (7) is not necessarily satisfied. Consequently, the valves in the network have to be adjusted such that Eq. (7) holds. Sibeijn et al. [17] prove that for each set of mass flows that satisfies Eq. (A.1), there exist configurations of valve openings and pump operations such that Eq. (7) is satisfied *if* a sufficient number of valves are placed across the network. Mathematically, at least n_f (number of free mass flows) valves must be placed such that the matrix $\boldsymbol{BV}$ is of full rank where $\boldsymbol{B}$ maps the valves to the respective network loops, and $\boldsymbol{V} \in \{0,1\}^{|\mathcal{E}|\times|\mathcal{E}_v|}$ indicates where valves are placed in the network [17].

$$V_{e,v} = \begin{cases} 1, & \text{if valve } v \text{ is on edge } e \\ 0, & \text{otherwise} \end{cases} \tag{A.2}$$

This requirement reflects the physical limitation that a valve cannot simultaneously balance the pressure in more than one loop. In addition, having redundant valves may be necessary to ensure that the mathematical solution is physically valid, as valves can only increase, and not reduce, pressure drops. Further details can be found in [17].

A.1.3. Thermal transients of the heat carrier

In large DHNs, it can take hours for the hot water front from the producer to reach the consumer. The time delays vary depending on the mass flow in the network, which complicates the tracking of hot water fronts [49]. The "node method" (see e.g. [12]) is commonly used to simulate thermal transients in DHNs. However, this method relies on integer variables to track historical time delays, making it unsuitable for optimization. This limitation can be addressed by applying iterative algorithms to solve the optimization [12] or by precalculating the integer variables [8] to simplify the optimization.

Alternatively, thermal transients can be simulated using other methods, such as finite-difference methods [50]. For instance, the first-order implicit upwind scheme in Eq. (15) is a viable option. While finite-difference methods do not require integer variables, they are nonlinear and therefore still challenging to optimize.

A.1.4. Constant mass flows or temperatures

Even when applying the model simplifications presented in the previous sections, the optimization problem may still be intractable. However, when either mass flows or temperatures are fixed, the optimization problem can be significantly simplified, and in the case of constant mass flows, even reduced to a linear problem. Both formulations, fixed mass flows and fixed temperatures, are substantiated by being common operation modes in practice [24,29].

Constant-flow formulation

If mass flows are constant, the mass flows and flow directions across the thermal network can be precalculated offline and used as parameters. This avoids challenge (C1) because flow directions are known, challenge (C2) because Kirchhoff's loop law need not be evaluated during optimization, and challenges (C3) and (C4) because the nonlinearity is simplified to linearity. Advantageously, a constant-flow formulation allows the heat carrier to function as thermal storage by varying its temperature. Disadvantageously, a constant-flow formulation may result in high return temperatures [51].

Constant-temperature formulations

In constant-temperature formulations, the supply temperature is held constant, which also motivates neglecting the thermal transients of the heat carrier. Instead, the temperature at the end of the pipeline can be estimated by assuming steady-state as

$$T^{L}_{\mathrm{p},k} = T_{\mathrm{amb},k} + (T^{0}_{\mathrm{p},k} - T_{\mathrm{amb},k}) e^{-L_{\mathrm{p}} U_{\mathrm{p}} d_{\mathrm{p}} \pi / c \dot{m}_{\mathrm{p},k}} \tag{A.3}$$

where $T^{L}_{\mathrm{p},k}, T^{0}_{\mathrm{p},k}$, are the outlet and inlet temperatures of the pipe at time step k, $T_{\mathrm{amb},k}$ denotes the ambient air temperature and $L_{\mathrm{p}}, U_{\mathrm{p}}, d_{\mathrm{p}}$ are the length, heat transfer coefficient, and diameter of the pipe, and $c, \dot{m}_{\mathrm{p},k}$ represent the heat carriers heat capacity and mass flow.

Although assuming steady-state temperatures reduces the complexity compared to computing temperature dynamics, the optimization problem remains nonlinear.

Consequently, the optimization model is often further simplified by neglecting heat losses [51] or by assuming constant temperature losses [27]. With these assumptions, all temperatures in the supply part of the network are fixed. Finally, it is possible to fix all temperatures across the DHN by additionally assuming constant temperature drops across the consumers, a simplification that can be supported by field measurements [51].

In summary, while a constant-temperature formulation simplifies challenges (C3) and (C4) by fixing temperatures or simplifying temperature computations, challenges (C1) and (C2), related to non-convexity and unknown flow directions, still require appropriate model simplifications. As a result, constant-temperature formulations typically still involve nonlinearity and may involve integer variables.

A.2. Comparison of different model simplifications

Five combinations of model simplifications are evaluated to determine their feasibility and suitability for MPC. The models differ in the following aspects:

- Some models assume constant mass flows or supply temperatures.
- Some models assume a constant temperature drop across consumers.
- The thermal model of the heat carrier is either transient (Eq. (15)), steady-state (Eq. (A.3)), or assumes no heat losses, leading to constant temperatures in the supply and return pipes.

The first-order implicit upwind scheme (Eq. (15)) becomes inaccurate for sample times larger than $\Delta t = 900\,\mathrm{s}$. Consequently, models including thermal transients use a sample time of $\Delta t = 900\,\mathrm{s}$ and the other models adopt $\Delta t = 3600\,\mathrm{s}$. Table A.7 presents the five models (M1 to M5).

Remark 3 (*Model Classes*)**.** Although the optimization problem includes only continuous variables, internal reformulations of the nonlinear constraints introduce integer variables, yielding a mixed-integer nonlinear program (MINLP) for the steady-state case. When fixing the temperatures throughout the DHN by neglecting heat losses, the model class reduces to a quadratically-constrained program (QCP), as the nonlinearities in the thermal model disappear. The model (M3) is convex since the non-convex bilinear equations become linear, and the mixing equation is no longer required.

For all configurations M1 to M5, flow directions are assumed to be known a priori, which holds true for tree networks and is a simplification for meshed networks. The convex reformulation of Kirchhoff's loop law (Appendix A.1.2) is applied to all model formulations with variable flows.

The optimizations are performed with CMPC on a Laptop equipped with an Intel Core i7-1370P CPU@1.90 GHz and 64 GB RAM. The optimization is aborted if it does not improve for two minutes, and the optimality gap from Gurobi is shown in Table A.8. Problems without a feasible solution after 15 min are considered intractable. While Table A.8 presents the results for a single exemplary time step, tests across different time steps confirm that these values are representative.

The results presented in Table A.8 show a considerable difference between convex and non-convex models. Convex models (M1 and M3) consistently achieve solution times within a few seconds and exhibit zero optimality gap across all tested scenarios. In contrast, non-convex models (M2, M4, and M5) demonstrate severely limited scalability. While M2 remains solvable, its computation time increases significantly with the number of flexible buildings, and the optimality gap remains substantial (up to 35.6%) as the solver stalls. Models M4 and M5, which incorporate transient thermal dynamics and free mass flows, become intractable under all tested conditions.

Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.energy.2026.140746.

Data availability

The utilized data are shared as supplementary material.

Table A.7
Overview of the model formulations.

Name	Model type	Mass flows	Supply temp.	Temp. drop	Thermal model	Sample time	Convex	Model class
M1	CF-VT	Constant	Free	Free	Transient	900 s	✓	LP
M2	VF-CT	Free	Constant	Constant	Steady-state	3600 s	✗	MINLP
M3		Free	Constant	Constant	Constant	3600 s	✓	QCP
M4		Free	Constant	Constant	Transient	900 s	✗	MIQCP
M5	VF-VT	Free	Free	Free	Transient	900 s	✗	MIQCP

Table A.8
Computation time and optimality gap for one time step with different numbers of flexible buildings.

	Computation time and optimality gap with J flexible buildings					
	$J = 0$		$J = 60$		$J = 120$	
	Comp. time	Opt. gap	Comp. time	Opt. gap	Comp. time	Opt. gap
M1	1 s	0%	7 s	0	13 s	0%
M2	90 s	0%	400 s	35.6%	568 s	31.5%
M3	0 s	0%	0.2 s	0%	1.2 s	0%
M4	intractable		intractable		intractable	
M5	intractable		intractable		intractable	

References

[1] Lund H, Østergaard PA, Nielsen TB, Werner S, Thorsen JE, Gudmundsson O, Arabkoohsar A, Mathiesen BV. Perspectives on fourth and fifth generation district heating. Energy 2021;227:120520. http://dx.doi.org/10.1016/j.energy.2021.120520.

[2] David A, Mathiesen BV, Averfalk H, Werner S, Lund H. Heat roadmap europe: Large-scale electric heat pumps in district heating systems. Energies 2017;10(4):578. http://dx.doi.org/10.3390/en10040578.

[3] Guelpa E, Verda V. Demand response and other demand side management techniques for district heating: A review. Energy 2021;219:119440. http://dx.doi.org/10.1016/j.energy.2020.119440.

[4] Vandermeulen A, Reynders G, van der Heijde B, Vanhoudt D, Salenbien R, Saelens D, Helsen L. Sources of energy exibility in district heating networks: Building thermal inertia versus thermal energy storage in the network pipes. In: Proceedings of USim conference 2018: 1st USim conference of IBPSA-Scotland, vol. 1, Glasgow, Scotland: IBPSA Scotland; 2018, p. 114–21.

[5] Vandermeulen A, van der Heijde B, Helsen L. Controlling district heating and cooling networks to unlock flexibility: A review. Energy 2018;151:103–15. http://dx.doi.org/10.1016/j.energy.2018.03.034.

[6] Kaspar K, Ouf M, Eicker U. A critical review of control schemes for demand-side energy management of building clusters. Energy Build 2022;257:111731. http://dx.doi.org/10.1016/j.enbuild.2021.111731.

[7] Cai H, You S, Wu J. Agent-based distributed demand response in district heating systems. Appl Energy 2020;262:114403. http://dx.doi.org/10.1016/j.apenergy.2019.114403.

[8] Maurer J, Illerhaus J, Soneira PJ, Hohmann S. Distributed optimization of district heating networks using optimality condition decomposition. Energies 2022;15(18):6605. http://dx.doi.org/10.3390/en15186605.

[9] Qin Q, Gosselin L. Community-based transactive energy market concept for 5th generation district heating and cooling through distributed optimization. Appl Energy 2024;371:123666. http://dx.doi.org/10.1016/j.apenergy.2024.123666.

[10] Taylor M, Marjanovic O, Parisio A. Decentralized supervisory control of networked multienergy buildings. IEEE Trans Control Syst Technol 2024;32(3):945–59. http://dx.doi.org/10.1109/TCST.2023.3346829.

[11] Jansen J, Jorissen F, Helsen L. Optimal control of a fourth generation district heating network using an integrated non-linear model predictive controller. Appl Therm Eng 2023;223:120030. http://dx.doi.org/10.1016/j.applthermaleng.2023.120030.

[12] Li Z, Wu W, Shahidehpour M, Wang J, Zhang B. Combined heat and power dispatch considering pipeline energy storage of district heating network. IEEE Trans Sustain Energy 2016;7:12–22. http://dx.doi.org/10.1109/TSTE.2015.2467383.

[13] Lefebure N, Khosravi M, Hudoba de Badyn M, Bünning F, Lygeros J, Jones C, Smith RS. Distributed model predictive control of buildings and energy hubs. Energy Build 2022;259:111806. http://dx.doi.org/10.1016/j.enbuild.2021.111806.

[14] Zheng W, Lu H, Zhang M, Wu Q, Hou Y, Zhu J. Distributed energy management of multi-entity integrated electricity and heat systems: A review of architectures, optimization algorithms, and prospects. IEEE Trans Smart Grid 2024;15(2):1544–61. http://dx.doi.org/10.1109/TSG.2023.3310947.

[15] Krug R, Mehrmann V, Schmidt M. Nonlinear optimization of district heating networks. Optim Eng 2021;22(2):783–819. http://dx.doi.org/10.1007/s11081-020-09549-0.

[16] de Giuli LB, Bella AL, Scattolini R. Physics-informed neural network modeling and predictive control of district heating systems. IEEE Trans Control Syst Technol 2024. http://dx.doi.org/10.1109/TCST.2024.3355476.

[17] Sibeijn M, Ahmed S, Khosravi M, Keviczky T. Economic nonlinear model predictive control of prosumer district heating networks. IEEE Trans Control Syst Technol 2025;1–16. http://dx.doi.org/10.1109/TCST.2025.3561501.

[18] La Bella A, Nigro L, Scattolini R. Predictive control and benefit sharing in multi-energy systems. IEEE Trans Control Syst Technol 2024;32(2):368–83. http://dx.doi.org/10.1109/TCST.2023.3310891.

[19] Hering D, Cansev ME, Tamassia E, Xhonneux A, Müller D. Temperature control of a low-temperature district heating network with model predictive control and mixed-integer quadratically constrained programming. Energy 2021;224:120140. http://dx.doi.org/10.1016/j.energy.2021.120140.

[20] Saletti C, Gambarotta A, Morini M. Development, analysis and application of a predictive controller to a small-scale district heating system. Appl Therm Eng 2020;165:114558. http://dx.doi.org/10.1016/j.applthermaleng.2019.114558.

[21] Quaggiotto D, Vivian J, Zarrella A. Management of a district heating network using model predictive control with and without thermal storage. Optim Eng 2021;22(3):1897–919. http://dx.doi.org/10.1007/s11081-021-09644-w.

[22] Schlichting H, Gersten K. Grenzschicht-Theorie. Berlin/Heidelberg: Springer-Verlag; 2006, http://dx.doi.org/10.1007/3-540-32985-4.

[23] van der Heijde B, Fuchs M, Ribas Tugores C, Schweiger G, Sartor K, Basciotti D, Müller D, Nytsch-Geusen C, Wetter M, Helsen L. Dynamic equation-based thermo-hydraulic pipe model for district heating and cooling systems. Energy Convers Manage 2017;151:158–69. http://dx.doi.org/10.1016/j.enconman.2017.08.072.

[24] Zheng J, Zhou Z, Zhao J, Wang J. Effects of the operation regulation modes of district heating system on an integrated heat and power dispatch system for wind power integration. Appl Energy 2018;230:1126–39. http://dx.doi.org/10.1016/j.apenergy.2018.09.077.

[25] Langner F, Wang W, Frahm M, Hagenmeyer V. Model predictive control of distributed energy resources in residential buildings considering forecast uncertainties. Energy Build 2024;303:113753. http://dx.doi.org/10.1016/j.enbuild.2023.113753.

[26] Harb H, Boyanov N, Hernandez L, Streblow R, Müller D. Development and validation of grey-box models for forecasting the thermal response of occupied buildings. Energy Build 2016;117:199–207. http://dx.doi.org/10.1016/j.enbuild.2016.02.021.

[27] Frölke L, Sousa T, Pinson P. A network-aware market mechanism for decentralized district heating systems. Appl Energy 2022;306:117956. http://dx.doi.org/10.1016/j.apenergy.2021.117956.

[28] Bhattacharya S, Chandan V, Arya V, Kar K. Demand response for thermal fairness in district heating networks. IEEE Trans Sustain Energy 2019;10:865–75. http://dx.doi.org/10.1109/TSTE.2018.2852629.

[29] Zhou Y, Shahidehpour M, Wei Z, Li Z, Sun G, Chen S. Distributionally robust unit commitment in coordinated electricity and district heating networks. IEEE Trans Power Syst 2020;35(3):2155–66. http://dx.doi.org/10.1109/TPWRS.2019.2950987.

[30] Korpela T, Kaivosoja J, Majanne Y, Laakkonen L, Nurmoranta M, Vilkko M. Utilization of district heating networks to provide flexibility in CHP production. Energy Procedia 2017;116:310–9. http://dx.doi.org/10.1016/j.egypro.2017.05.077.

[31] Boyd S. Distributed optimization and statistical learning via the alternating direction method of multipliers. Found Trends® Mach Learn 2010;3(1):1–122. http://dx.doi.org/10.1561/2200000016.

[32] IEEE standard for floating-point arithmetic. 2019, p. 1–84. http://dx.doi.org/10.1109/IEEESTD.2019.8766229, IEEE Std 754-2019 (Revision IEEE 754-2008).
[33] European Comission. PVGIS. 2025, https://re.jrc.ec.europa.eu/pvg_tools. [Accessed 02 April 2025].
[34] Bundesnetzagentur. SMARD strommarktdaten. 2025, URL https://www.smard.de/home/downloadcenter/download-marktdaten. [Accessed 04 July 2025].
[35] Hagenmeyer V, Kemal Çakmak H, Düpmeier C, Faulwasser T, Isele J, Keller HB, Kohlhepp P, Kühnapfel U, Stucky U, Waczowicz S, Mikut R. Information and communication technology in energy lab 2.0: Smart energies system simulation and control center with an open-street-map-based power flow simulation example. Energy Technol 2016;4(1):145–62. http://dx.doi.org/10.1002/ente.201500304.
[36] Wiegel F, Wachter J, Kyesswa M, Mikut R, Waczowicz S, Hagenmeyer V. Smart energy system control laboratory – a fully-automated and user-oriented research infrastructure for controlling and operating smart energy systems. At - Autom 2022;70(12):1116–33. http://dx.doi.org/10.1515/auto-2022-0018.
[37] Lee ZE, Zhang KM. Generalized reinforcement learning for building control using behavioral cloning. Appl Energy 2021;304:117602. http://dx.doi.org/10.1016/j.apenergy.2021.117602.
[38] Ueno T, Meier A. A method to generate heating and cooling schedules based on data from connected thermostats. Energy Build 2020;228:110423. http://dx.doi.org/10.1016/j.enbuild.2020.110423.
[39] Andersson JAE, Gillis J, Horn G, Rawlings JB, Diehl M. CasADi – A software framework for nonlinear optimization and optimal control. Math Program Comput 2019;11(1):1–36. http://dx.doi.org/10.1007/s12532-018-0139-4.
[40] Gurobi Optimization, LLC. Gurobi optimizer reference manual. 2024, URL https://www.gurobi.com.
[41] Nigro L, La Bella A, Casella F, Scattolini R. Control-oriented modeling, simulation, and predictive control of district heating networks. IEEE Trans Autom Sci Eng 2025;22:7064–79. http://dx.doi.org/10.1109/TASE.2024.3460173.
[42] Dagdougui H, Bagheri F, Le H, Dessaint L. Neural network model for short-term and very-short-term load forecasting in district buildings. Energy Build 2019;203:109408. http://dx.doi.org/10.1016/j.enbuild.2019.109408.
[43] Eseye AT, Lehtonen M. Short-term forecasting of heat demand of buildings for efficient and optimal energy management based on integrated machine learning models. IEEE Trans Ind Inform 2020;16(12):7743–55. http://dx.doi.org/10.1109/TII.2020.2970165.
[44] Idowu S, Saguna S, Åhlund C, Schelén O. Applied machine learning: Forecasting heat load in district heating system. Energy Build 2016;133:478–88. http://dx.doi.org/10.1016/j.enbuild.2016.09.068.
[45] Runge J, Saloux E. A comparison of prediction and forecasting artificial intelligence models to estimate the future energy demand in a district heating system. Energy 2023;269:126661. http://dx.doi.org/10.1016/j.energy.2023.126661.
[46] Hua P, Wang H, Xie Z, Lahdelma R. District heating load patterns and short-term forecasting for buildings and city level. Energy 2024;289:129866. http://dx.doi.org/10.1016/j.energy.2023.129866.
[47] Langner F, Frahm M, Wang W, Matthes J, Hagenmeyer V. Hierarchical-stochastic model predictive control for a grid-interactive multi-zone residential building with distributed energy resources. J Build Eng 2024;89:109401. http://dx.doi.org/10.1016/j.jobe.2024.109401.
[48] Hohmann M, Warrington J, Lygeros J. A two-stage polynomial approach to stochastic optimization of district heating networks. Sustain Energy Grids Netw 2019;17:100177. http://dx.doi.org/10.1016/j.segan.2018.11.003.
[49] Duquette J, Rowe A, Wild P. Thermal performance of a steady state physical pipe model for simulating district heating grids with variable flow. Appl Energy 2016;178:383–93. http://dx.doi.org/10.1016/j.apenergy.2016.06.092.
[50] Zheng X, Shi K, Wang Y, You S, Zhang H, Zhu C, Li L, Wei S, Wang N. Performance analysis of three iteration-free numerical methods for fast and accurate simulation of thermal dynamics in district heating pipeline. Appl Therm Eng 2020;178:115622. http://dx.doi.org/10.1016/j.applthermaleng.2020.115622.
[51] Cai H, Ziras C, You S, Li R, Honoré K, Bindner HW. Demand side management in urban district heating networks. Appl Energy 2018;230:506–18. http://dx.doi.org/10.1016/j.apenergy.2018.08.105.