Efficient Optimization of Inertial Energy Networks Using Physics-Informed Spatio-Temporal Model

Taha Boussaid
taha.boussaid@insa-lyon.fr
INSA Lyon, CNRS, CETHIL, UMR 5008 & LIRIS, UMR 5205
Villeurbanne, France

François Rousset
francois.rousset@insa-lyon.fr
INSA Lyon, CNRS, CETHIL, UMR 5008
Villeurbanne, France

Marc Clausse
marc.clausse@insa-lyon.fr
INSA Lyon, CNRS, CETHIL, UMR 5008
Villeurbanne, France

Vasile-Marian Scuturici
marian.scuturici@insa-lyon.fr
INSA Lyon, CNRS, LIRIS, UMR 5205
Villeurbanne, France

Abstract

Mitigating climate change calls for a transition to more sustainable energy systems. One of the key levers is the efficient deployment of district energy networks which integrate diverse low-carbon sources. Inertial energy networks, characterized by their large-scale infrastructure, are an example of systems designed to provide heating and cooling to various consumers. However, the complexity of their dynamics such as multi-timescale responses, nonlinear behaviors and intermittence of renewable sources result in high computational cost for accurate simulation and optimization. To overcome this, we introduce a system-agnostic modeling framework that leverages the spatio-temporal structure of these systems to develop an efficient and scalable physics-informed state-space surrogate model. Coupled with model-based and learning-based optimizers, the framework enables full dynamic optimization and faster decision-making while drastically reducing computing time. Benchmarks against rule-based and physics-based strategies demonstrate that our approach achieves competitive energy cost savings while cutting the runtime from over a month to just few hours.

Keywords

Spatio-temporal model, predictive control, sustainable energy systems, reinforcement learning.

ACM Reference Format:
Taha Boussaid, François Rousset, Marc Clausse, and Vasile-Marian Scuturici. 2025. Efficient Optimization of Inertial Energy Networks Using Physics-Informed Spatio-Temporal Model. In *Proceedings of Temporal Graph Learning Workshop, SIGKDD International Conference on Knowledge Discovery and Data Mining 2025 (TGL Workshop, KDD 2025).* ACM, New York, NY, USA, 6 pages.

TGL Workshop, KDD 2025, Toronto, Canada

ACM ISBN 978-1-4503-XXXX-X/2018/06

1 Introduction

In recent reports, the international energy agency (IEA) outlines the urgent need to increase the efficient deployment of inertial energy networks with multiple low-carbon footprint energy sources to reach net-zero emissions by 2050 [14]. In fact, nearly half of the world energy demand in buildings was attributed to heating, with fossil fuels still accounting for 60% of the heat production in 2023 [15].

The term *inertial* refers to the system's thermal inertia, where the thermal mass of pipes and storage introduces delays between heat production and distribution. Multiple sources, such as biomass boilers, geothermal, solar, and gas (as backup), can operate simultaneously, while thermal storage enables asynchronous and flexible operation [17]. Each source has distinct, often nonlinear dynamics and constraints. For example, biomass units respond slowly, while solar output is weather-dependent. The combination of diverse sources, operational constraints, and stochastic demand makes accurate modeling and optimization computationally intensive [2, 10, 18, 20].

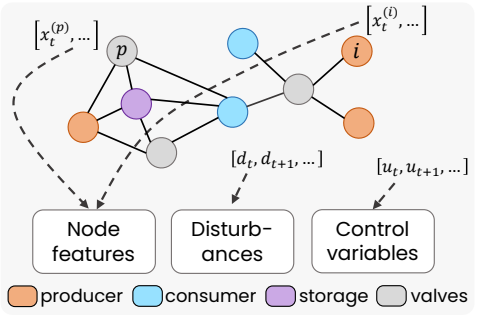


Figure 1: Graph representation of an inertial energy network: nodes represent producers, consumers, and other physical entities, while edges denote pipes transporting the heat carrier. Node-level features and graph-level disturbance and control variables are also shown.

To address the computational bottleneck, several approaches have been explored in the literature, including linearization techniques [23, 29] and reduced order modeling [12, 25]. Nevertheless, these approaches often lead to a simplification of the dynamics

and reduces their fidelity in system representation. In contrast, deep learning-based approaches have recently gained attention due to their ability to model complex nonlinear behaviors while ensuring reduced inference time. Such methods have covered a wide range of applications, including weather and thermal load forecasting, fluid dynamics, and energy systems control via deep reinforcement [5, 11, 26, 27].

For energy networks, recent studies proposed the use of deep learning-based surrogate models to reduce simulation time [7, 8, 24]. Given their inherent topology, these networks are suited to graph-based representations, as illustrated in Fig. 1, making graph neural networks (GNNs) a relevant approach for surrogate modeling [6, 19, 32]. However, no systematic methodology or unified design framework has been developed for inertial energy networks. Besides, prior works often focus the optimization on short periods or specific seasons, and restrict their analysis to single-producer configurations. Furthermore, spatio-temporal formulations, which would better capture the interactions between dynamic temporal behaviors and spatial dependencies remain underexplored.

Contributions: This work introduces a comprehensive and scalable framework for dynamic simulation and control optimization of inertial energy networks. Our approach leverages the inherent graph topology of these systems to construct a spatio-temporal graph neural network. In contrast to prior work, we explicitly incorporate a generic hydraulic mass conservation equation, resulting in a physics-informed surrogate model with improved reliability. Moreover, the method seamlessly handle various energy sources at different locations of the network along with multiple consumer nodes. The development phase incorporates an adaptation of Gaussian scaling and window slicing to augment historical time-series data and expose the surrogate model to physically plausible scenarios during the training. Both simulation and optimization are framed within a state-space formalism, facilitating the integration with a variety of model-based and learning-based optimization algorithms. Finally, the impact and relevance of our approach is demonstrated on a real-world system over a full year of operation. It achieves competitive energy cost reductions while being up to three orders of magnitude faster than conventional physics-based simulations. The full methodology is schematized in Fig. 2

2 Related work

Predictive control, as schematized in Fig. 2, requires a system model to perform predictive simulations and accurately capture its response to control signals. In control theory, this dynamical model is often expressed through an ordinary differential equation (ODE) in a state-space [4]. Mathematically, the control is performed for an optimization time horizon, $\mathcal{H}^{\text{opt}} = [0, t_f]$, and can be formalized as:

$$\begin{aligned} \frac{dx(t)}{dt} &= f(x(t),\ u(t),\ d(t)),\ \text{and } x(0) = x_0, \\ C\left(t_f,\ u\right) &= \int_0^{t_f} g(t,\ x_u(t),\ u(t))\,dt + h(t_f, x_u(t_f)), \\ \text{s.t. } x_u(t) &= \text{argmin } C(t_f, u) \text{ and } l(t,\ x(t),\ u(t)) \le 0\ \forall t, \end{aligned} \tag{1}$$

where f denotes the nonlinear system dynamics, with $x \in \mathbb{R}^{n_x}$ the state vector, $u \in \mathbb{R}^{n_u}$ the control inputs, and $d \in \mathbb{R}^{n_d}$ the external disturbances. The cost function C comprises a running cost g and a terminal cost h evaluated at the end of the optimization horizon $\mathcal{H}^{\text{opt}}$, subject to state and control constraints l.

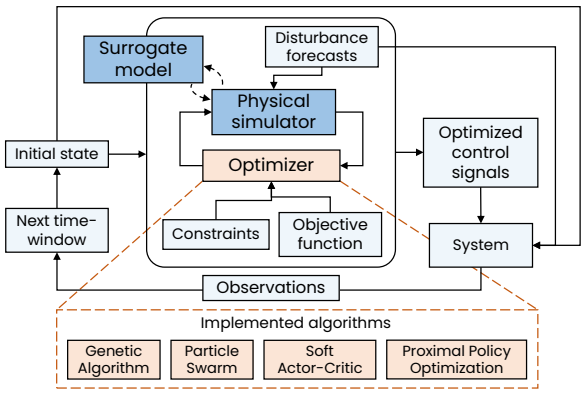


Figure 2: Methodological framework for system-agnostic simulation and control optimization of inertial energy systems. The physical-simulator is replaced by a surrogate model, that can be then coupled to different optimizers.

State-space models can be learned in two distinct ways, discrete-time (DT) or continuous-time (CT) models. The latter usually requires initial state estimation while DT models are more common and easier to construct as data is numerically represented via discrete elements [1, 3]. For inertial energy networks, surrogate DT models have been proposed in [6, 8, 9, 21]. For instance, the authors of [8] associated a recurrent neural network (RNN) to each consumer node in a district heating network (DHN), but only considered a single-producer case and relied on manually connecting RNN cells, approach where spatio-temporal GNNs would be more suitable. In [21], a long-short term memory (LSTM) surrogate model was coupled with a deep reinforcement learning agent (soft-actor critic) to enable efficient district energy system management, reducing the energy costs by almost 3% for a three months period. Nevertheless, a yearly assessment is essential for capturing seasonal variations and demand patterns. We address these gaps by introducing a spatio-temporal graph convolutional network that captures both temporal dynamics and spatial dependencies, critical for inertial energy networks where fast hydraulic responses interact with slow thermal transients. The architecture adopts a *time-then-space* formulation, shown to be more expressive than joint *time-and-space* models [13]. Our physics-informed, system-agnostic model delivers reliable predictions and integrates seamlessly with various optimizers, as demonstrated in a full-year real-world case.

3 Methodology

3.1 State-Space Surrogate Model

The network is represented as a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, with $|\mathcal{V}|$ nodes and $|\mathcal{E}|$ edges. Its state is defined through the evolution of state variables defined at each node (e.g., x = mass flow rate). Moreover the entire system is influenced by exogenous variables, classified as disturbances (e.g., d = weather) or control variables (e.g., u = power). The optimal control formulated in Eq. 1 seeks to optimize the network trajectory to minimize a cost function C while satisfying a set

of constraints over $\mathcal{H}^{\text{opt}}$. We propose to reformulate this problem such that future system evolution is inferred as follows:

$$x_+^{\mathcal{H}^{\text{sm}}} = f_\theta\left(x_-^{\mathcal{H}^{\text{sm}}},\ u_+^{\mathcal{H}^{\text{sm}}},\ d_+^{\mathcal{H}^{\text{sm}}}\right), \tag{2}$$

where $\mathcal{H}^{\text{sm}}$ is the predictive range of the surrogate model. The subscript + indicates predicted variables, meaning values from time t to time $t + \mathcal{H}^{\text{sm}}$. Subscript − indicates past observations or measurements of state variables from $t - \mathcal{H}^{\text{sm}}$ to t. In Eq. 2, the learned state-space model (f_θ) predicts future system states based on past states, future control variables and forecasted disturbances. The inference function f_θ results from the proposed architecture shown in Fig. 3, where in addition to past variables, disturbance forecasts and future control signals are broadcasted to each node.

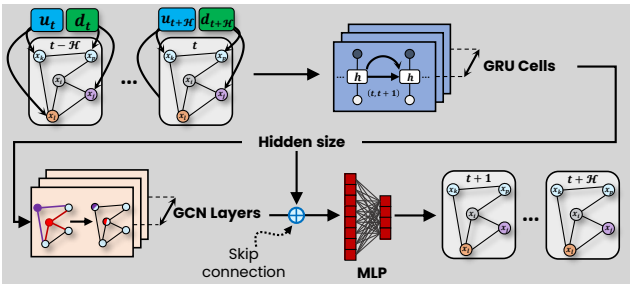


Figure 3: State-space surrogate model predicting future state trajectory $x_+^{\mathcal{H}}$ from past states $x_-^{\mathcal{H}}$, future controls $u_+^{\mathcal{H}}$, and disturbance forecasts $d_+^{\mathcal{H}}$.

Next, these combined node features go through a encoder-processor-decoder architecture. Gated recurrent units (GRU) are used to encode temporal dependencies, the processing is performed using stacked graph convolutions (GCN) and the decoding is made through a MLP. The resulting hypeparameters: number of GRU, GCN, the hidden size, batch size and learning rate are optimized using the asynchronous successive halving algorithm from [16]. The surrogate model is trained to minimize the following loss function:

$$\begin{aligned}
\mathcal{L}_\theta &= \frac{1}{N_b}\sum_b \frac{1}{\mathcal{H}}\sum_{t_b}^{t_b+\mathcal{H}}\left[\frac{1}{\mathcal{V}}\sum_n \left\|\hat{x}_{b,n,t} - x_{b,n,t}\right\|_2^2 + \lambda\cdot\mathcal{F}_m^2\left(\hat{x}\right)\right], \\
\text{s.t. } \mathcal{F}_m\left(\hat{x}\right) &= \sum_{\text{producers}} \hat{x}_{b,n,t} \;-\; \sum_{\text{consumers}} \hat{x}_{b,n,t}\ , \\
\hat{x}_+^{\mathcal{H}} &= f_\theta\left(x_-^{\mathcal{H}},\ u_+^{\mathcal{H}},\ d_+^{\mathcal{H}}\right).
\end{aligned} \tag{3}$$

The loss function is weighted with a physical constraint term represented by $\mathcal{F}_m$ (through λ, a hyperparameter). It represents the mass conservation over the network: the sum of the mass flow rates received by the consumers must be equal to the sum of mass flow rates sent by the producers. The loss is averaged and calculated over a batch of size N_b and across all the nodes in the network $\mathcal{V}$. This surrogate model is benchmarked against a standard model, vector autoregressive (VARx), then a multi-layer perceptron (MLP) and a recurrent neural network based on gated recurrent units (RNN).

3.2 Augmentation and Training

Conventional physics-based approaches rely on a digital twin of the real system to assess consistency and system response under forecasted disturbances (e.g., weather, demand). This high-fidelity simulator is typically validated using real-world measurements prior to deployment. Here, we propose to leverage real-world historical measurements of both control and disturbances to augment the training dataset using this physical simulator as illustrated in Appendix B. We propose combining Gaussian scaling and window slicing from [28] to augment simulation inputs. Let $z_t = [z_t, \ldots, z_{t+k}]$ denotes a timeserie spanning k time-steps, its augmented version is computed as follows:

$$\mathcal{T}\left[z_t\right] = [\underbrace{\gamma_1 \cdot z_t, \ldots,}_{\Delta\ window}\ \gamma_2 \cdot z_{t+\Delta}, \ldots, \underbrace{\gamma_j \cdot z_{t+j\Delta}, \ldots, \gamma_j \cdot z_{t+k}}_{\Delta\ window}], \tag{4}$$

where $\gamma \sim \mathcal{N}(1, \sigma_{aug}^2)$, $\mathcal{N}$ being the normal distribution and $1 < \Delta < k$ is a fixed time window. This parameter acts as the sampling frequency for scaling coefficients, i.e., the data is scaled every Δ steps. Here, only control and disturbance variables are subject to augmentation. State variables cannot be independently augmented, as their evolution is intrinsically governed by system dynamics and constrained by physical laws. By augmenting only inputs (control and disturbances), the resulting state trajectories are augmented and remain physically consistent, as they are generated through simulation using the underlying physical model.

After optimizing the hyperparameters, the best model configuration is trained for optimal performance using AdamW. All experiments are conducted on a 48 GB NVIDIA A40 GPU. The benchmarked models are configured with the same parameter count obtained from the hyperparameter optimization of the spatio-temporal model. The train/validation/test split is illustrated in Appendix B.

3.3 Optimizers

Predictive control relies on the surrogate model for optimization. Evolutionary algorithms such as genetic algorithms (GA) and particle swarm optimization (PSO), widely used in the literature [22, 31], are implemented in this work. Additionally, we explore deep reinforcement learning, which has recently shown strong performance in similar settings [11, 21], using both on-policy (proximal policy optimization, PPO) and off-policy (soft actor-critic, SAC) agents.

4 Experiments and Results

4.1 Experimental Setup

The surrogate model is validated on a real-world system integrating three energy sources, biomass, gas, and solar, along with thermal storage. A graph-based representation is shown in Fig. 4, and the corresponding physical schematic and details are given in Appendix A. This system was selected due to the availability of a full year of historical control, weather, and demand data, as well as access to a high-fidelity physical simulator. For scalability assessment, and given the lack of open-access datasets, we generated synthetic data from various network topologies using an open-source physical solver [30], with examples provided in Appendix D.

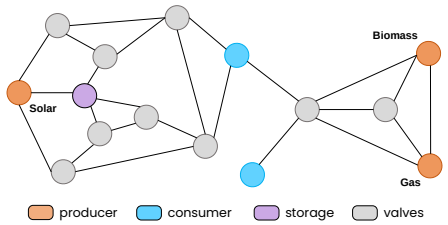


Figure 4: Graph representation of a real-world system with biomass and gas boilers, a solar field, and thermal storage. Larger topologies with more consumer nodes are tested and shown in Appendix D.

4.2 Performance Analysis

Two accuracy metrics are used: mean squared error (MSE) to assess the absolute error on normalized state variables (mass flow rates, temperatures, and valve positions), and symmetric mean absolute percentage error (sMAPE) to express relative error. Precision is measured using the coefficient of determination (R^2). Results are shown in Fig. 5. The proposed PI-STGCN model outperforms all

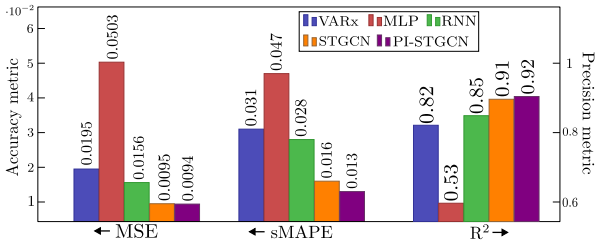


Figure 5: Performance comparison of the physics-informed spatio-temporal graph convolution (PI-STGCN) with other models.

others, achieving a sMAPE of approximately 1.3% and a global R^2 of 0.92. The integration of the mass constraint ($\mathcal{F}_m$) in the loss function improves performance, as seen in the comparison between STGCN (trained to minimize MSE only) and PI-STGCN. The MLP model performs notably worse due to its lack of both temporal and spatial encoding, which are captured by the other models. Examples of timeseries predictions are given in Appendix E. Finally, the scalability assessment results are reported in Table 1. The surrogate model maintains high accuracy across all networks, with R^2 above 0.97 and MSE decreasing as network size increases, indicating robustness. Physical simulations become more expensive, with runtime rising from 27s to 665s, while surrogate inference remains efficient, increasing from 20ms to 330ms, an average reduction of three orders of magnitude. On a per-node basis, inference time decreases with network size (from 5ms to 3.3ms), while simulation time remains stable at 6.7s per node, showing that inference scales sub-linearly with network size.

4.3 Optimization Results

After validating the learned state-space model, we apply it to optimal control. A full-year optimization is run using a sliding window

Table 1: Scalability evaluation on test cases from Appendix D.

Network	$\mathcal{V}=4$	$\mathcal{V}=6$	$\mathcal{V}=14$	$\mathcal{V}=100$
MSE ($\times 10^{-3}$)	1.2	1.2	0.19	0.08
sMAPE	0.21	0.18	0.15	0.10
R^2	0.98	0.99	0.97	0.98
Simulation/sample (s)	27	41	203	665
Inference/sample (ms)	20	23	52	330
Inference/node (ms)	5	3.8	3.7	3.3

of $\mathcal{H}^{\text{opt}} = 3$ days. The objective is to minimize running fuel cost (biomass and gas), with constraints ensuring temperature stays above a comfort threshold and control variations remain limited. GA and PSO were stopped when cost variations reached a 5-monetary units tolerance for 10 generations or after 200 iterations. PPO and SAC were trained until achieving the maximum cumulative reward (defined in Appendix C), requiring 1.5 million (2h 40min) and 1 million (5h 55min) training iterations, respectively. Additionally, we include a comparison with the GA optimizer directly coupled to the physical simulator through a functional mock-up unit (FMU). Fuel cost reduction results are summarized in Table. 2 and the computational time of each method is provided in Fig. 6.

Table 2: Energy cost reduction (%) by month for each optimization method: functional mock-up (FMU), genetic algorithm (GA), particle swarm optimization (PSO), soft actor-critic (SAC) and proximal policy optimization (PPO).

Month	1	2	3	4	5	6	7	8	9	10	11	12
FMU	0.5	2.7	0.6	0.8	2.1	1.4	0.3	1.1	4.7	2.5	0.3	0.3
GA	0.2	2.0	0.1	0.6	1.9	1.8	0.2	0.8	0.5	2.3	0.2	0.1
PSO	0	1.6	0	0	0.2	0.1	0	0	0	0.8	0	0
SAC	0	1.1	0	0	0.6	0	2.2	1.7	0	0	0	0
PPO	0	0.1	0	0	0	0	4.7	1.4	0	0	0	0

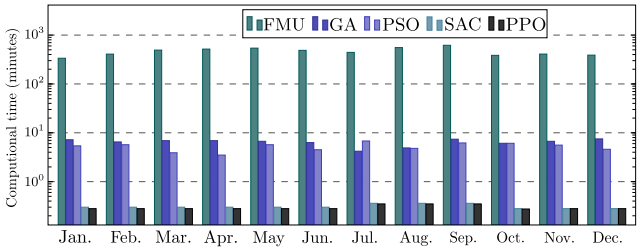


Figure 6: Average runtime required for the optimization of $\mathcal{H}^{\text{opt}}$ period across the month and for each optimizer.

The results show that fuel cost reductions remain below 5% across all methods, reflecting the problem's inherent difficulty, especially in winter, when high heat demand limits system flexibility. Among all methods, physics-based GA yields the best fuel savings in most months (up to 4.73% in September) but requires 31.2 days to run. This is attributed to the high accuracy of the physical simulator's predictions, which enhances optimization precision. While

there is a slight reduction in optimization quality due to prediction errors, the surrogate-based GA offers a strong trade-off, securing the second-best results in most months and even outperforming the FMU-based GA in June (1.82% reduction), while being considerably faster with a runtime of just 10.2 hours. PSO is faster but less effective, often stuck in local optima. RL methods show inconsistent performance and frequent constraint violations, despite PPO achieving peak savings in July.

5 Conclusion

We proposed a hybrid optimization and simulation framework for district inertial energy networks. The methodology combines an efficient spatio-temporal surrogate model with heuristic and reinforcement learning. The surrogate-based GA achieved the best trade-off, reducing computation time from 31 days to 10.2 hours while maintaining strong performance. Results also highlighted seasonal control limitations, particularly in winter. RL methods, though capable of reaching optimization goals, frequently violated constraints, this suggests a reformulation of the reward function or the use of a hierarchical agents approach. Future work will further analyze and refine the control strategies learned by each method.

References

[1] Ibrahim Ayed, Emmanuel de Bézenac, Arthur Pajot, Julien Brajard, and Patrick Gallinari. 2019. Learning dynamical systems from partial observations. *arXiv preprint arXiv:1902.11136* (2019).

[2] Gregor Becker, Christian Klemm, and Peter Vennemann. 2022. Open source district heating modeling tools—A comparative study. *energies* 15, 21 (2022), 8277.

[3] Gerben I. Beintema, Maarten Schoukens, and Roland Tóth. 2023. Continuous-time identification of dynamic state-space models by deep subspace encoding. In *The Eleventh International Conference on Learning Representations*.

[4] Pierre Clément Blaud, Philippe Chevrel, Fabien Claveau, Pierrick Haurant, and Anthony Mouraud. 2023. From multi-physics models to neural network for predictive control synthesis. *Optimal Control Applications and Methods* 44, 3 (2023), 1394–1411.

[5] Diego Botache, Jens Decke, Winfried Ripken, Abhinay Dornipati, Franz Götz-Hahn, Mohamed Ayeb, and Bernhard Sick. 2024. Enhancing multi-objective optimisation through machine learning-supported multiphysics simulation. In *Joint ECML-PKDD*. Springer, 297–312.

[6] Taha Boussaid, François Rousset, Vasile-Marian Scuturici, and Marc Clausse. 2024. Enabling fast prediction of district heating networks transients via a physics-guided graph neural network. *Applied Energy* 370 (2024), 123634.

[7] Sam J Cox, Dongsu Kim, Heejin Cho, and Pedro Mago. 2019. Real time optimal control of district cooling system with thermal energy storage using neural networks. *Applied energy* 238 (2019), 466–480.

[8] Laura Boca de Giuli, Alessio La Bella, and Riccardo Scattolini. 2024. Physics-informed neural network modeling and predictive control of district heating systems. *IEEE Transactions on Control Systems Technology* (2024).

[9] Steven de Jongh, Sina Steinle, Anna Hlawatsch, Felicitas Mueller, Michael Suriyah, and Thomas Leibfried. 2021. Neural Predictive Control for the Optimization of Smart Grid Flexibility Schedules. In *2021 56th International Universities Power Engineering Conference (UPEC)*. IEEE, 1–6.

[10] Régis Delubac, Sylvain Serra, Sabine Sochard, and Jean-Michel Reneaume. 2021. A dynamic optimization tool to size and operate solar thermal district heating networks production plants. *Energies* 14, 23 (2021), 8003.

[11] Jifei Deng, Miro Eklund, Seppo Sierla, Jouni Savolainen, Hannu Niemistö, Tommi Karhela, and Valeriy Vyatkin. 2023. Deep reinforcement learning for fuel cost optimization in district heating. *Sustainable Cities and Society* 99 (2023), 104955.

[12] Basak Falay, Gerald Schweiger, Keith O'Donovan, and Ingo Leusbrock. 2020. Enabling large-scale dynamic simulations and reducing model complexity of district heating and cooling systems by aggregation. *Energy* 209 (2020), 118410.

[13] Jianfei Gao and Bruno Ribeiro. 2022. On the equivalence between temporal and static graph representations for observational predictions. In *International Conference on Machine Learning*. PMLR.

[14] IEA. 2023. Net Zero Roadmap: A Global Pathway to Keep the 1.5 °C Goal in Reach. Licence: CC BY 4.0.

[15] IEA. 2024. *IEA, Heating*. Technical Report. International Energy Agency. https://www.iea.org/energy-system/buildings/heating

[16] Lisha Li, Kevin Jamieson, Afshin Rostamizadeh, Katya Gonina, Moritz Hardt, Benjamin Recht, and Ameet Talwalkar. 2018. Massively parallel hyperparameter tuning. (2018).

[17] Henrik Lund, Poul Alberg Østergaard, Miguel Chang, Sven Werner, Svend Svendsen, Peter Sorknæs, Jan Eric Thorsen, Frede Hvelplund, Bent Ole Gram Mortensen, Brian Vad Mathiesen, et al. 2018. The status of 4th generation district heating: Research and results. *Energy* 164 (2018), 147–159.

[18] Jona Maurer, Jochen Illerhaus, Pol Jané Soneira, and Sören Hohmann. 2022. Distributed optimization of district heating networks using optimality condition decomposition. *Energies* 15, 18 (2022), 6605.

[19] Damian Owerko, Fernando Gama, and Alejandro Ribeiro. 2020. Optimal power flow using graph neural networks. In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 5930–5934.

[20] Matija Pavičević, Tomislav Novosel, Tomislav Pukšec, and Neven Duić. 2017. Hourly optimization and sizing of district heating systems considering building refurbishment–Case study for the city of Zagreb. *Energy* 137 (2017), 1264–1276.

[21] Giuseppe Pinto, Davide Deltetto, and Alfonso Capozzoli. 2021. Data-driven district energy management with surrogate models and deep reinforcement learning. *Applied Energy* 304 (2021), 117642.

[22] Chao Qin, Qingyou Yan, and Gang He. 2019. Integrated energy systems planning with electricity, heat and gas using particle swarm optimization. *Energy* 188 (2019), 116044.

[23] Jim Rojer, Femke Janssen, Thijs van der Klauw, and Jacobus van Rooyen. 2024. Integral techno-economic design & operational optimization for district heating networks with a Mixed Integer Linear Programming strategy. *Energy* 308 (2024), 132710.

[24] Etienne Saloux, Jason Runge, and Kun Zhang. 2023. Operation optimization of multi-boiler district heating systems using artificial intelligence-based model predictive control: Field demonstrations. *Energy* 285 (2023), 129524.

[25] Johan Simonsson, Khalid Tourkey Atta, and Wolfgang Birk. 2022. Reduced-Order Modeling of Thermal Dynamics in District Energy Networks using Spectral Clustering. In *2022 IEEE Conference on Control Technology and Applications (CCTA)*. IEEE, 144–150.

[26] Yogesh Verma, Markus Heinonen, and Vikas Garg. 2024. ClimODE: Climate and Weather Forecasting with Physics-informed Neural ODEs. In *The Twelfth International Conference on Learning Representations*.

[27] Zhijin Wang, Xiufeng Liu, Yaohui Huang, Peisong Zhang, and Yonggang Fu. 2023. A multivariate time series graph neural network for district heat load forecasting. *Energy* 278 (2023), 127911.

[28] Qingsong Wen, Liang Sun, Fan Yang, Xiaomin Song, Jingkun Gao, Xue Wang, and Huan Xu. 2021. Time Series Data Augmentation for Deep Learning: A Survey. In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21*. 4653–4660. doi:10.24963/ijcai.2021/631

[29] Marco Wirtz, Lisa Neumaier, Peter Remmen, and Dirk Müller. 2021. Temperature control in 5th generation district heating and cooling networks: An MILP-based operation optimization. *Applied Energy* 288 (2021), 116608.

[30] Francesco Witte and Ilja Tuschy. 2020. TESPy: Thermal Engineering Systems in Python. *Journal of Open Source Software* 5, 49 (2020), 2178. doi:10.21105/joss.02178

[31] Min Wu, Pengcheng Du, Meihui Jiang, Hui Hwang Goh, Hongyu Zhu, Dongdong Zhang, and Thomas Wu. 2022. An integrated energy system optimization strategy based on particle swarm optimization algorithm. *Energy Reports* 8 (2022), 679–691.

[32] Yin Yu, Xinyuan Jiang, Daning Huang, Yan Li, Meng Yue, and Tianqiao Zhao. 2024. PIDGeuN: Graph neural network-enabled transient dynamics prediction of networked microgrids through full-field measurement. *IEEE Access* (2024).

Appendices

A Real-world system

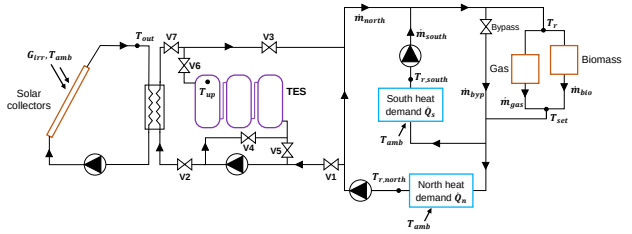


Figure 7: Schematic diagram of the real-world system. The contour colors correspond to the ones used in the graph representation.

Several valves can be seen between the solar field and the storage, they allow different cycles: charging or discharging the storage, or direct injection from the solar field to the network. In the state-space formulation, mass flow rates are state variables for boilers and control valves. Boiler mass flow rates are proportional to power output, while valve flow rates indicate open or closed states. For consumers and storage, fluid temperature is considered as the state variable. In storage, temperature reflects stored heat, while consumer temperatures indicate the heat demand.

The system also depends on external disturbances: solar irradiance G_{irr} and ambient temperature T_{amb}, which affect solar energy production, as well as heat demand in the northern $\dot{Q}_n$ and southern $\dot{Q}_s$ clusters.

B Augmentation

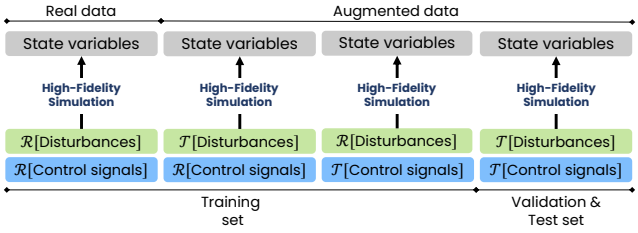


Figure 8: Augmentation process: The original dataset, derived from real-world exogenous variables ($\mathcal{R}$), is used as a baseline. Each exogenous variable is then independently augmented ($\mathcal{T}$) and simulated using a high-fidelity simulator.

C Reward function

The cost function C is set to the economic operational cost of the network and can be written as:

$$C\left(t_f,\ u\right) = \int_0^{t_f} c_{\text{bio}} \cdot \dot{Q}_{\text{bio}}(t) + c_{\text{gas}} \cdot \dot{Q}_{\text{gas}}(t)\ dt. \tag{5}$$

Where c_{bio} and c_{gas} are fuel costs (in M.U/kWh) and $\dot{Q}$ the energy provided by biomass and gas respectively. Moreover, the optimal control must satisfy two constraints:

$$\left(\frac{u_{i,t} - u_{i,t-1}}{\Delta t}\right)^2 \leq \delta_u, \text{ and } x_{i,t} \geq x_{min}\ \forall t \in \mathcal{H}^{opt},\ i \in \{\text{n}, \text{s}\}. \tag{6}$$

The first constraint limits control variable variations to protect hydraulic pumps, while the second ensures outlet temperatures stay above a threshold for comfort and safety.

Constrained optimizations must be adapted to align with the reward-learning framework of RL agents. Accordingly, both PPO and SAC were trained using the following reward function:

$$r_t = -\frac{C_t}{C^*} - \frac{1}{\mathcal{H}} \sum_i \sum_t^{t+\mathcal{H}} \left[\lambda_u \left(\frac{u_{i,t} - u_{i,t-1}}{\Delta t}\right)^2 + \lambda_x e^{x_{min} - x_{i,t}}\right]. \tag{7}$$

This formulation balances three terms: minimizing economic costs (normalized by the rule-based cost C^*), enforcing smooth control variations (weighted by λ_u), and ensuring temperature constraints (weighted by λ_x).

D Scalability cases

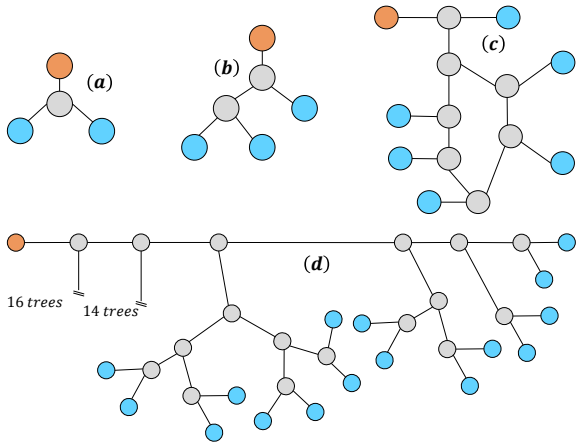
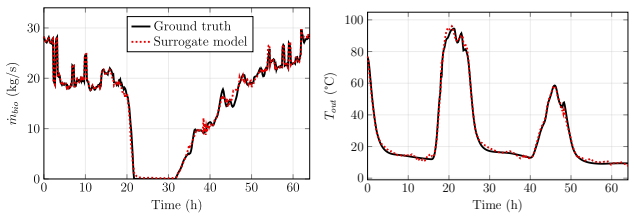


Figure 9: Synthetic use cases for surrogate model scalability assessment. The number of nodes is 4 in (a), 6 in (b), 14 in (c) and 100 in (d).

E Dynamics Learning



(a) Biomass flow rate **(b) Solar field temperature**

Figure 10: Examples of PI-STGCN predictions (red dotted) vs. high-fidelity simulation (black) over an illustrative period, capturing both fast (a) and slow (b) dynamics.