

## Article

# Investigation of Predictive Regulation Strategy of Secondary Loop in District Heating Systems

Zhongbo Li <sup>1</sup>, Zheng Luo <sup>2</sup>, Ning Zhang <sup>2</sup>, Xiaojie Lin <sup>1,3,\*</sup> , Wei Huang <sup>1</sup>, Encheng Feng <sup>1</sup> and Wei Zhong <sup>1,4,5</sup><sup>1</sup> College of Energy Engineering, Zhejiang University, Hangzhou 310027, China<sup>2</sup> Polytechnic Institute, Zhejiang University, Hangzhou 310015, China<sup>3</sup> Jiaxing Research Institute, Zhejiang University, Jiaxing 314499, China<sup>4</sup> Zhejiang Key Laboratory of Clean Energy and Carbon Neutrality, Zhejiang University, Hangzhou 310027, China<sup>5</sup> Shanghai Institute for Advanced Study, Zhejiang University, Shanghai 201203, China

\* Correspondence: xiaojie.lin@zju.edu.cn; Tel.: +86-0571-8795-1058

**Abstract:** The urban energy system is greatly dependent on the District Heating System (DHS). However, many difficulties with regulation and control are caused by its large scale and numerous coupling variables. Additionally, reliance on manual experience means it can be challenging to guarantee heating comfort and effectiveness in the regulation of DHS. This paper proposes a data-driven temperature response prediction model to predict secondary loop supply temperature based on the heating substation's historical operating status, valve opening degree, weather conditions, etc. Further, the XGBoost model was established in this article with different input and prediction steps. The results show that the XGBoost model with 72 input steps and 24 prediction steps has better performance. As an application example, the model was applied to an urban central heating system. Based on this data-driven model, different operation strategies on primary loop valve opening are compared for temperature response analysis. Operators can check the temperature responses of different valve control strategies before being applied. This paper guides the regulation behavior of the DHS, which is of great significance for the operation of the actual DHS.

**Keywords:** district heating system; predictive regulation strategy; XGBoost; valve control strategies



**Citation:** Li, Z.; Luo, Z.; Zhang, N.; Lin, X.; Huang, W.; Feng, E.; Zhong, W. Investigation of Predictive Regulation Strategy of Secondary Loop in District Heating Systems. *Sustainability* **2023**, *15*, 3524. <https://doi.org/10.3390/su15043524>

Academic Editor: Graziano Salvalai

Received: 2 December 2022

Revised: 17 January 2023

Accepted: 13 February 2023

Published: 14 February 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The DHS, an essential part of the energy system in a city, has been widely adopted in northern China. At the end of 2021, the national urban central heating area reached 10,603,000,000 m<sup>2</sup>, up 7.30% from the previous year. It is noticeable that the consumption of DHS occupying building energy consumption is more than 50%. Increasing buildings' energy efficiency has enormous potential for energy savings. The control strategy is the main factor restricting the heating system's efficiency. The valve on the primary loop is adapted to adjust the temperature of secondary supply water, as shown in Figure 1.

The heating substations are located between the primary and secondary loops, where the transfer of heat between the primary and secondary loops is accomplished. If running well, the heating substations supply water with the expected temperature, which means the indoor temperatures of all apartments are comfortable. However, with the influence of the lag effect [1] in the heat transfer process and nonlinearity between variables, reaching the target temperature on the secondary loop is especially difficult. In practice, to reduce user complaints, the regulators may continuously supply heating, leading to overheating in buildings. This causes excessive heat loss, which can account for 10–20% of the total [2].

Existing studies about DHS mainly focused on the primary loop, but there is a lack of research on the secondary loop of DHS. Zheng et al. [3] developed a scheduling model of a thermos-electric integrated energy system considering DHS thermal inertia, in which a complete hydraulic and thermal model is integrated to realize the dynamic temperature

calculation of the whole network. Wang et al. [4] presented a method for pressure regulation in DHS. Genetic algorithms are used to optimize the distributed variable speed pump control technique, and a real DHS system is used to test the effectiveness of the strategy. Stevanovic et al. [5] presented dynamic external conditions (wind intensity and solar radiation intensity) and dynamic demand adjustment methods of heat source and pumping station in DHS. Gu et al. [6] proposed a hybrid control scheme by applying electric control valves with distributed variable speed pumps, which shows a great advantage in reducing the pressure of the pipeline networks and effectively weakened the hydraulic imbalance. Gustafsson et al. [7] found that it is possible and advantageous to utilize the primary supply temperature for radiator system control while maintaining comfort and no additional temperature sensors were needed. Gregor et al. [8] analyzed the causes of degradation of temperature difference between supply and return flow and quantitatively evaluated the numerous possible causes of degradation using a dynamic simulation environment. Developing commissioning tools and approaches to rapidly detect faults in hydraulic networks is recommended.

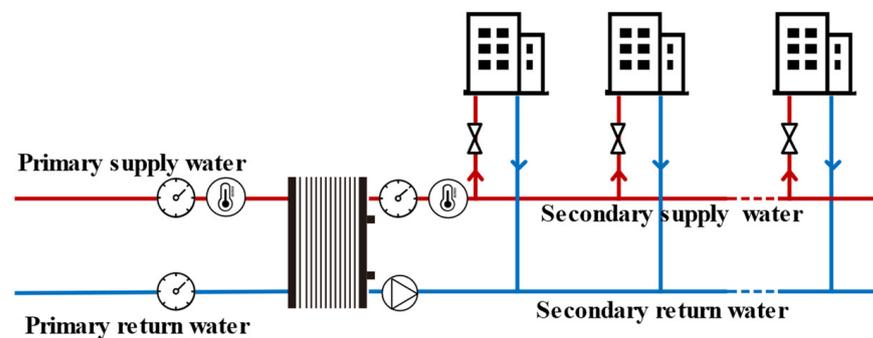


Figure 1. Structure of heating substation.

The modeling of the heating system is the basis of the regulation of DHS. Most of the literature focuses on generating control strategies with mechanism-based models, whose modeling process is relatively complex. Data-driven models are mostly used for key parameter prediction, which is less used in the studies of generating regulation strategy of DHS. Mendes et al. [9] established a mechanism-based model applied to build thermal analysis and control systems design. A lumped approach was used to model the room air temperature and a multi-layer model for the building envelope. Karlsson [10] described a conceptual model for investigating the effects of increasing the thermal storage capacity of building materials. The effects of wall thickness, wall area, free solar radiation, and other factors on building heat storage were studied. Zhun [11] developed a building energy demand forecasting model based on the decision tree method, which can classify and predict categorical variables. Its competitive advantage over other widely used modeling techniques, such as the regression and ANN methods, lies in generating accurate predictive models with interpretable flowchart-like tree structures that enable users to extract useful information quickly. Touzani [12] proposed an energy consumption baseline modeling method that utilizes a gradient boosting machine. The results show that the linear regression machine model  $R^2$  improved by 80% compared with the industry best practice model based on gradient boosting and the random forest algorithm. However, this study has not been validated in a real system and lacks engineering significance. Magnier [13] firstly used neural networks to describe the behavior of buildings and then combined neural networks with multi-objective genetic algorithms for optimization studies of thermal comfort and energy consumption in buildings. Machado et al. [14] addressed a comprehensive nonlinear ODE-based thermo-hydraulic model of a DHS. This study proposes a new hydraulic solution method, but the computational speed is not brought up. Xu et al. [15] developed an integrated model for simulating the thermal and hydraulic behavior of the heating system with various operation cases, and the results show that when the set value of the TRV is kept at 2~3, its effectiveness in reducing the overheating phenomena caused by

excessive water flow rate. Chicherin et al. [16] utilized the scaling design of heat demand and the weighted moving average to investigate variations and peak values of actual heat demand profiles in a district heating network. The results show that operational heat data are utilized to discover that almost no weather correlation exists during warm months when supply temperatures exceed 60 °C; thermal inertia of buildings affects their behavior differently in terms of needed space heating.

As for the research on the regulation of the secondary loop of the heating system, many models were proposed for efficient and low-carbon operation and control. Wang et al. [17] introduced a thermo-hydraulic couplings model built on TRNSYS 18, two stochastic parameters, infiltration rates, and increased thermal resistance of buried pipes were considered. However, they were lacking in model accuracy for engineering applications. Zheng et al. [18] introduced the intermittent heating mode for promoting wind power integration in an integrated heat and power dispatch system based on a real DHS featuring multiple heat sources and looped networks. Zhao et al. [19] proposed the indirect heating, direct heating, and water source heat pump auxiliary heating modes depending on different temperature levels. Sun et al. [20] put forward a control strategy of DHS that integrated characteristics of users' energy-saving behaviors and the combined control of feedforward and feedback were realized. Wang et al. [21] proposed a novel thermal energy flow model with transmission time delay and an optimal scheduling strategy for district-integrated heat and power systems. Cadau et al. [22] presented a model predictive control approach based on the prediction of the future evolution of the controlled system to manage district heating and cooling networks. Bojic et al. [23] proposed a steady-state, bottom-up approach, and sequential linear programming was used to solve the unbalanced distribution of heat in a DHS. Turski et al. [24] proposed the energetic effect of using buildings and a district heating network as thermal energy storage to compensate for the reduced heat output of the DHS. In Ruseljuk et al. [25] the adaptability of different CHP equipment for different district heating systems is compared, providing a reference for district heating system design planning. Garcia et al. [26] studied a district heating system connected to renewable energy. The related researches have been listed and classified in the Table 1.

**Table 1.** Summary of modeling and regulation in DHS.

Modeling Scale	Research Target	Highlights	Ref.
DHS	thermal characteristic of DHS	Model the room air temperature and a multi-layer model for the building envelope	[10]
	thermal characteristic of DHS	The effects of wall thickness, wall area, free solar radiation, and other factors on building heat storage were studied	[11]
	regulation of DHS	A building energy demand forecasting model based on the decision tree method	[12]
	regulation of DHS	A baseline modeling based on the gradient boosting machine to forecast the energy consumption.	[13]
DHS	thermal characteristic of DHS	Combine neural networks with multi-objective genetic algorithms for optimization studies of thermal comfort and energy consumption in buildings	[14]
	regulation of DHS	Address a comprehensive nonlinear ODE-based thermo-hydraulic model of the DHS	[15]
	thermal characteristic of DHS	Develop an integrated model for simulating the thermal and hydraulic behavior of the DHS	[16]
	thermal characteristic of DHS	Thermal inertia of buildings affects their behavior differently in terms of needed space heating	[17]
	regulation of DHS	A steady-state, bottom-up approach, and sequential linear programming was used to solve the unbalanced distribution of heat in a DHS	[25]
	regulation of DHS	The energetic effect of using buildings and a district heating network as thermal energy storage to compensate for the reduced heat output of the DHS	[26]

Table 1. Cont.

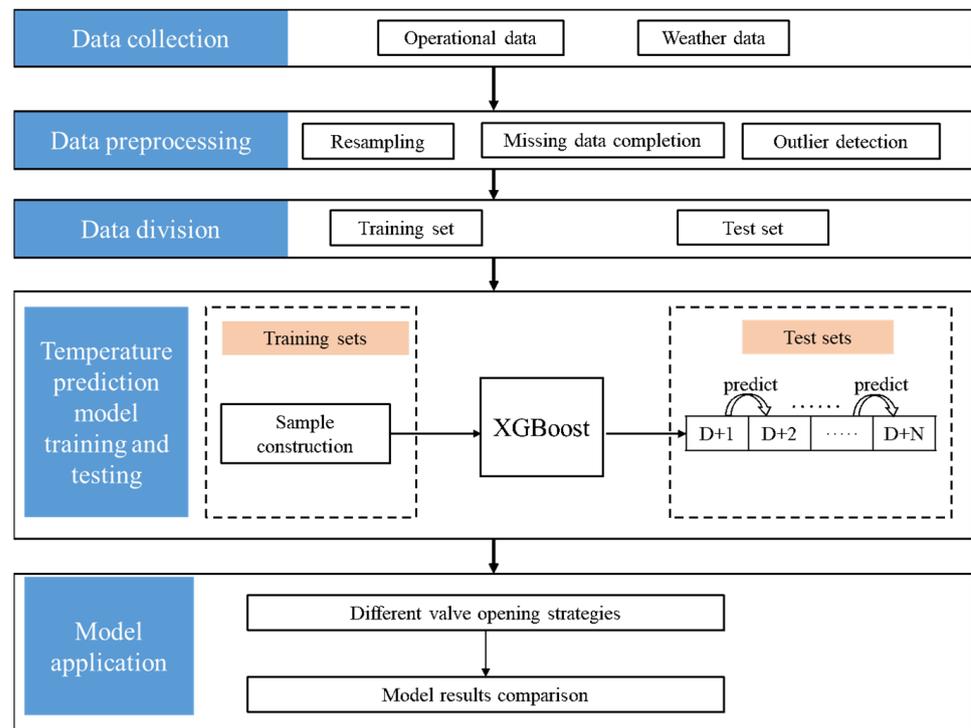
Modeling Scale	Research Target	Highlights	Ref.
Primary loop of DHS	regulation of DHS	Genetic algorithm is used to optimize the adjustment strategy of distributed variable speed pump in DHS.	[4]
	regulation of DHS	A demand regulation model of DHS heat source and pump station affected by external conditions	[5]
	regulation of DHS	A numerical model to predict the thermal transients in the DHS	[6]
	regulation of DHS	A hybrid control scheme with electric control valves with DVSPs to the DHS	[7]
	regulation of DHS	Primary supply temperature affects the result of the prediction for primary return temperature	[8]
	thermal characteristic of DHS	The reason for temperature difference degradation	[9]
Secondary loop of DHS	regulation of DHS	Study the effect of infiltration rates and increased thermal resistance of buried pipes to thermo-hydraulic couplings model	[18]
	regulation of DHS	Introduce the intermittent heating mode for promoting wind power integration in an integrated heat and power dispatch system	[19]
	regulation of DHS	Propose the indirect heating, direct heating, and water source heat pump auxiliary heating modes in the DHS	[20]
	regulation of DHS	A control strategy of DHS that realized combined control of feedforward and feedback	[21]
	regulation of DHS	A novel thermal energy flow model with transmission time delay in the DHS	[22]

Most current research concentrates on one single area of DHS regulation, and few studies offer a complete regulatory method for DHS regulation. To reduce energy consumption in the DHS, improve system efficiency, and reduce reliance on manual empirical decisions, in this paper, based on multiple features, a model to predict the demand of secondary loop supply temperature was established. Then, the mapping relationship between the secondary loop supply temperature and the primary loop valve opening is dug. The paper forms a complete DHS control solution that provides one-stop guidance for regulators.

## 2. Methodology

The research framework of this study is described in Figure 2, which includes five main processes. Firstly, a real DHS's raw data, which included climatic information, was gathered. Then, the raw data were subjected to data preprocessing in order to fill in gaps and address outliers. After that, the test set and the train set are divided in the ratio of 2:8. Then, the integration model was developed to forecast the secondary loop temperature response. Further, the evaluation method, root mean square error (RMSE), was used to analyze the models using the test dataset.

The extreme gradient boosting machine (XGBoost), which is developed by Chen [10], is a learning method based on a framework with gradient, which has been frequently used in prediction issues. The fundamental principle of the XGBoost aims to continuously calculate the current model's residuals by structuring a new model, and to serially cover all the models to obtain the final prediction.



**Figure 2.** Methodology framework of data processing and model deployment.

Data preprocessing is necessary before developing models because repeating data, missing values, and outliers are contained in the collected data caused by signal transmission and acquisition problems, which will decrease model performance. The data preprocessing is mainly divided into the three parts:

1. Remove repeating data:

A simultaneous storage of the raw data is possible due to reading or storage faults. As a result, only one piece of data must be retained rather than multiple.

2. Complete data and down-sampling:

Due to signal transmission issues or equipment, there could have been no data at certain time periods throughout the raw data collecting phase; thus, the values of missing data were replaced by average values.

3. Detect and replace outliers:

The outliers were found using the three-sigma approach, and then replaced with neighboring values.

### XGBOOST

The ensemble algorithm integrates multiple tree models to obtain a prediction model with relatively good performance. Its structure is shown in (1). The  $t$  is the number of tree models, and the  $f_K(x_i)$  is the  $K$ -th tree model.

$$\begin{aligned}
 \hat{y}_i^{(0)} &= 0 \\
 \hat{y}_i^{(1)} &= f_1(x_i) = \hat{y}_i^{(0)} + f_1(x_i) \\
 \hat{y}_i^{(2)} &= f_1(x_i) + f_2(x_i) = \hat{y}_i^{(1)} + f_2(x_i) \\
 &\dots \\
 \hat{y}_i^{(t)} &= \sum_{K=1}^t f_K(x_i) = \hat{y}_i^{(t-1)} + f_t(x_i)
 \end{aligned} \tag{1}$$

XGBoost is a typical ensemble algorithm with an outstanding performance with sparse data. Coincidentally, heating system data have the same characteristics. Its objective function is mainly composed of error terms and penalty terms, as shown in (2).

$$L(\phi) = \sum_i l(\hat{y}_i - y_i) + \sum_k \Omega(f_k) \quad (2)$$

$l(\hat{y}_i - y_i)$  is the error term, the  $i$  is represented as the  $i$ -th sample, which represents the error between the predicted value and true value of the heating system data.  $\Omega(f_k)$  is the penalty term,  $K$  is represented as  $K$ -th tree model. Additionally, the complexity of the model can be calculated by the penalty term. The smaller the complexity, the stronger the generalization ability. The specific penalty terms are shown in (3).

$$\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2 \quad (3)$$

$T$  is the number of leaf nodes and  $\omega_j$  is the weight of those. The minimum objective function computed is shown in (4).

$$\begin{aligned} Obj &= -\frac{1}{2} \sum_{j=1}^T \frac{G_j^2}{H_j + \lambda} + \gamma T \\ G_j &= \sum_{i \in I_j} g_i, H_j = \sum_{i \in I_j} h_i \\ g_i &= \partial_{\hat{y}^{(t-1)}} l(y_i, \hat{y}^{(t-1)}), h_i = \partial_{\hat{y}^{(t-1)}}^2 l(y_i, \hat{y}^{(t-1)}) \end{aligned} \quad (4)$$

The  $g_i$  and  $h_i$  is the first derivative and second derivative of the loss function. In this study, the squared loss function is applied for ease of calculation, as shown in (5).

$$l(y_i, \hat{y}^{(t-1)}) = (y_i - \hat{y}_i^{(t-1)})^2 \quad (5)$$

The first derivative and second derivative of the loss function can be calculated, as shown in (6) and (7).

$$g_i = \frac{\partial l(y_i, \hat{y}_i^{(t-1)})}{\partial \hat{y}_i^{(t-1)}} = -2(y_i - \hat{y}_i^{(t-1)}) \quad (6)$$

$$h_i = \frac{\partial^2 l(y_i, \hat{y}_i^{(t-1)})}{\partial (\hat{y}_i^{(t-1)})^2} = 2 \quad (7)$$

Substitute the above equation into the minimum objective function, and an approximation of the minimum objective function can be obtained, as shown in (8).

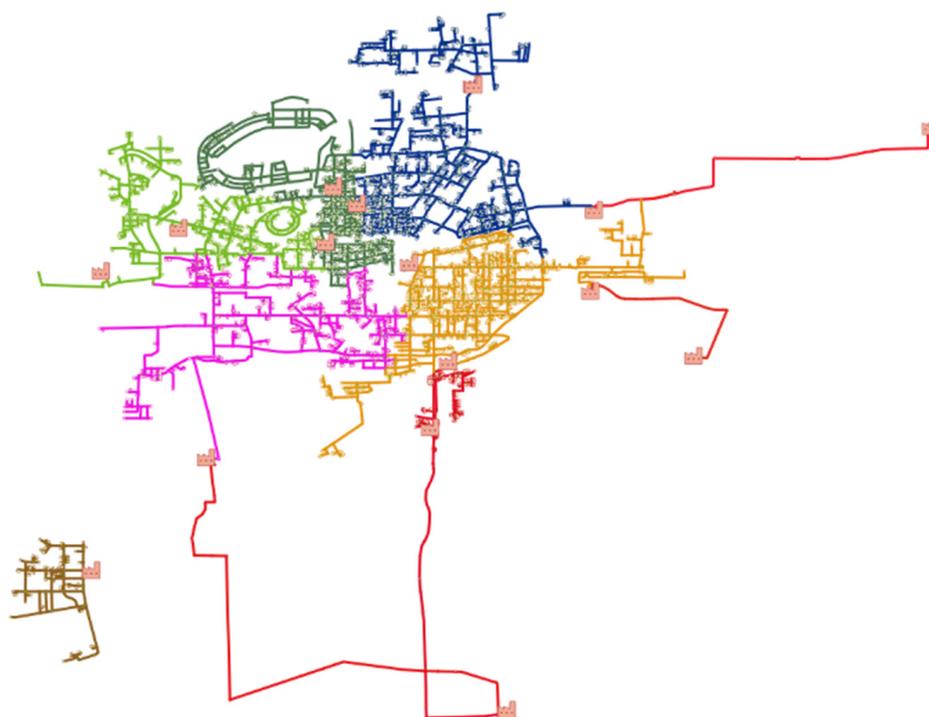
$$Obj^{(t)} \simeq \sum_{i=1}^n [l(y_i, \hat{y}^{(t-1)}) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)] + \Omega(f_t) + c \quad (8)$$

The tree added to the model has been determined, which means there are some gradient-independent terms in (8). The terms are removed from the minimum objective function, and (9) is finally obtained.

$$Obj^{(t)} \simeq \sum_{i=1}^n [g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)] + \Omega(f_t) \quad (9)$$

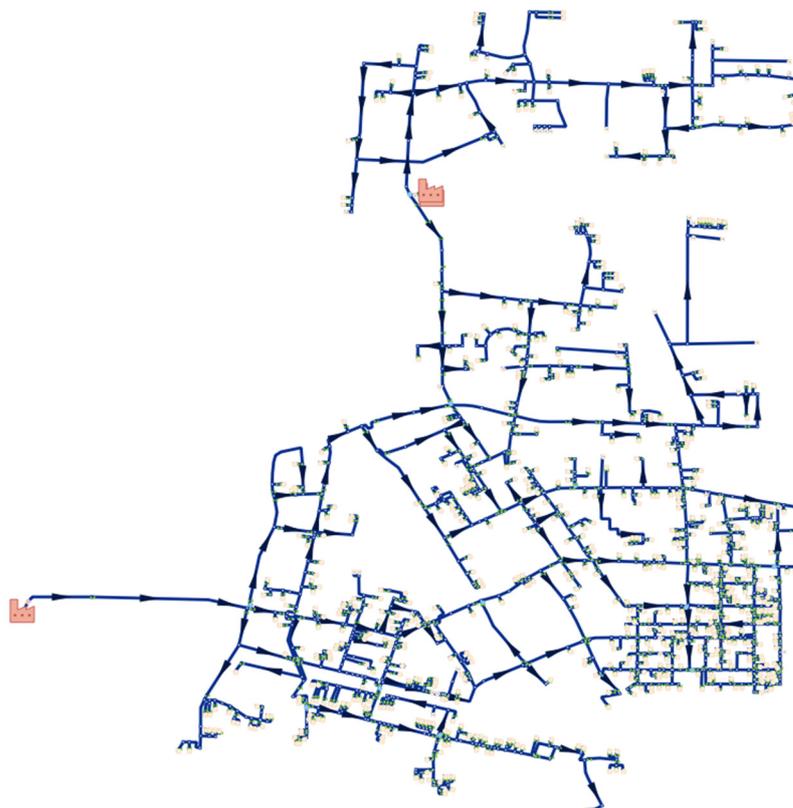
This paper adopts the sparse perception algorithm when constructing the tree model to handle the missing data or outliers in the heating system. When each node is established, a default direction is added so that the missing data samples are recognized and have no effect on model calculations.

The experimental data included in this study was obtained from a real DHS in the Chinese city of Zhengzhou. The case study in this paper is a cooperation project with Zhengzhou Heat Group Co., Ltd., Zhengzhou, China. In this DHS, there are 20 heat sources, 4 peripheral power plants, 6 pressure isolation stations, and 10 gas boiler houses for heating in different areas. Regarding the heat network, there are 10 independent operational areas in the whole network during the heating season. The whole area of this heat network reaches 136,948,500 km<sup>2</sup>, and the actual supply area of this heat network achieves 115,553,750 km<sup>2</sup>. For the heating substations that have been put into operation, the number of heating substations covering the whole network is about 2962 or so. The area is in the very high heating supply stage, with a source heat supply of 31.715 million GJ, of which 19.7 million GJ or 62% is supplied by cogeneration, gas boiler houses supply 12.015 million GJ or 38%, and the total consumption of natural gas is 355.686 million m<sup>3</sup>. As shown in Figure 3, there are 8 branches under Zhengzhou Heat Group Co., Ltd. to manage the heat network area, respectively, which is used as the research object of this paper. A public building is heated by the heating substation selected and its secondary loop. The details of the topology diagram structure in the northern direction of Figure 3 have been shown in Figure 4. In Figure 3, the different districts are represented by lines with different colors and the lines represent the pipework of primary loop in DHS. In Figure 4, a district pipework with two heat sources is indicated.



**Figure 3.** Topology of district heating network.

The data processing and modeling process in this paper is implemented on Python with jupyter. The hardware information are as shown in the following: 11th Gen Intel(R) Core(TM) i7-11800H, NVIDIA Geforce RTX 3060 GPU.



**Figure 4.** Part of the specific display in Figure 3.

In this study, raw data were collected from an actual system over a period of three months. The data interval was set at five minutes. To simplify the data, the raw data were down-sampled to an hourly frequency. This was achieved by calculating the average value every five minutes and using it as the representative value for the new hourly sample. The climate model strategy was also collected, which means the supply water temperature target value according to the outdoor air temperature and the climate model as shown in Table 2.

**Table 2.** Climate model of the experimental heating substation.

Air Temperature/ $^{\circ}\text{C}$	Heating Intensity per Unit Area/ $\text{W}\cdot\text{m}^{-2}$	Regulation Target Values		
		$\bar{T}_2/^{\circ}\text{C}$	$T_{r,1}/^{\circ}\text{C}$	$T_{s,2}/^{\circ}\text{C}$
−4	45.39	45.2	50.2	52.2
−3.5	45	45	50	52
−3	43.45	44.3	49.1	51
−2.5	42.48	43.8	48.6	50.4
−2	41.51	43.4	48	49.8
−1.5	40.54	42.9	47.4	49.2
−1	40.57	42.9	47.4	49.2
−0.5	39.6	42.5	46.9	48.6
0	38.63	42	46.3	48
0.5	37.66	41.5	45.7	47.4
1	36.69	41.1	45.2	46.8
1.5	35.72	40.6	44.6	46.2
2	34.75	40.1	44	45.5
2.5	33.78	39.7	43.4	44.9
3	32.81	39.2	42.8	44.3
3.5	30.84	38.2	41.6	43
4	29.97	37.7	41	42.3
4.5	28.9	37.2	40.4	41.7

Table 2. Cont.

Air Temperature/°C	Heating Intensity per Unit Area/W·m <sup>-2</sup>	Regulation Target Values		
		$\bar{T}_2/°C$	$T_{r,1}/°C$	$T_{s,2}/°C$
5	27.93	36.7	39.8	41.1
5.5	26.96	36.2	39.2	40.4
6	25.99	35.7	38.6	39.7
6.5	25.02	35.2	38	39.1
7	24.05	34.7	38	38.4
7.5	23.08	34.2	38	37.7
8	22.11	33.6	38	37.1
8.5	21.14	33.1	38	36.4
9	20.17	32.6	38	35.7
9.5	19.2	32	38	35
10	18.23	31.5	38	34.3

### 3. Results and Discussion

The predicted results of XGBoost are discussed in this section. Figures 5–7 show that the XGBoost model predicts the stepwise prediction error (RMSE) of 6, 12, and 24 steps in the future with different input step sizes. Table 3 lists the average error results of the model under different input and output step sizes.

Table 3. Average prediction error of XGBoost under varying input and output steps.

Input Steps	Prediction Steps		
	6	12	24
6	0.162	—	—
12	0.147	0.151	—
24	0.129	0.140	0.141
48	0.123	0.130	0.128
72	0.116	0.114	0.117

Judging from the stepwise prediction errors shown in Figures 5–7, the XGBoost model’s prediction error will reach the maximum at prediction steps 3 to 6, and the stepwise prediction error does not show an obvious upward trend as the step size increases.

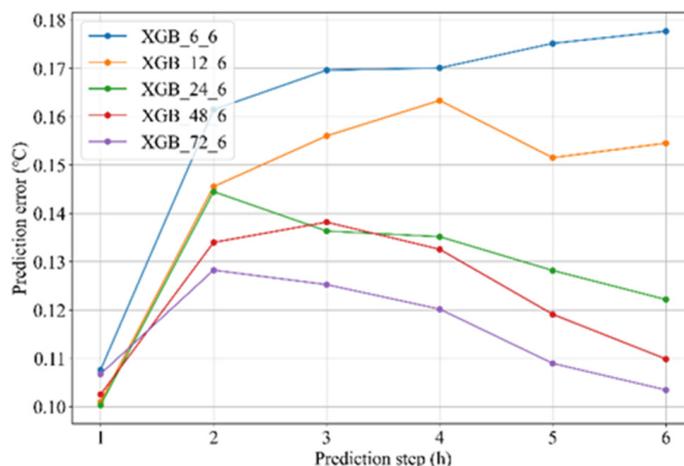
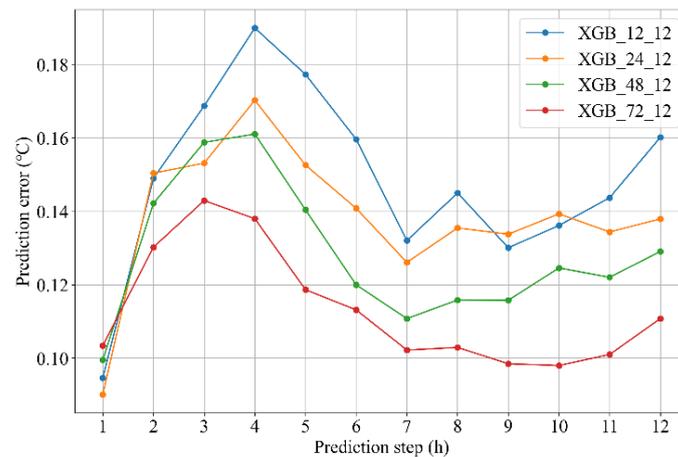
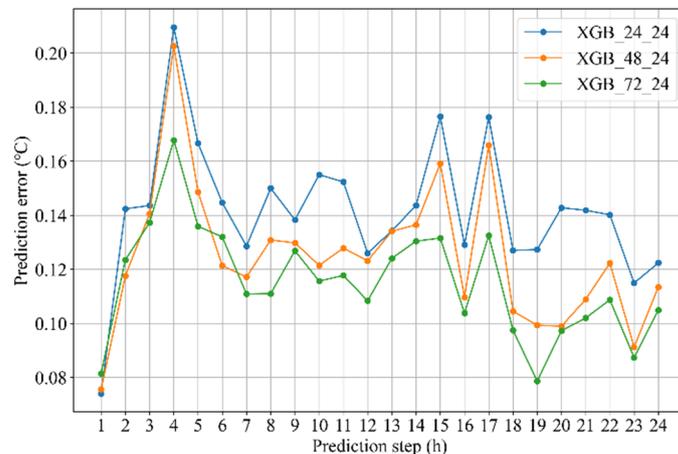


Figure 5. Prediction error in next 6 h under different input steps of XGBoost.



**Figure 6.** Prediction error in next 12 h under different input steps of XGBoost.



**Figure 7.** Prediction error in next 24 h under different input steps of XGBoost.

From the perspective of the prediction step size, the different XGBoost models do not show an obvious upward trend with the increase in the prediction steps. As it can be seen in Table 3, there is no obvious difference in the prediction error of the models under different input step sizes and prediction step sizes.

As Table 3 shows, the average prediction error of the XGBoost model decreases with the input step size increasing, when viewed from the perspective of the input step size. Take the prediction step size of 6 as an example: the average prediction errors when the input step sizes are 6, 12, 24, 48, and 72 reach 0.162, 0.147, 0.129, 0.123, and 0.116, respectively. The same goes for the prediction step size of 12 or 24. That means that for the scenario where the XGBoost model is used to forecast the temperature of the secondary loop supply water, the longer the input steps included in the samples (within 72 steps), the higher the accuracy of the prediction. The reason for this result is that the XGBoost model performs well on information extraction through the integration of multiple predictors. When the input step is longer, which means that more historical information is included, XGBoost can effectively extract more helpful information to improve prediction performance.

As shown in Table 3, the XGBoost models presented the lowest prediction error with an input step size of 72 and prediction step size of 24, which was adopted as the temperature response prediction model. The prediction takes control data of the experimental heating substation on 4 January 2020, as an application example. The data collection interval is down-sampled from five minutes to one hour.

The temperature of the secondary supply and return water is controlled by the primary loop valve. A predictive regulation model of heating substations with machine learning is established to solve the problems of regulation in the current urban central heating system,

combined with the actual heating substation operation data and weather data. In addition, based on the experience of regulators and the information on climate, the climate model was summed up. The climate model represents the quantitative relationship between the outdoor temperature and valve opening, as shown in Table 2. For example: when the external temperature is  $-3\text{ }^{\circ}\text{C}$ , the heating intensity should be  $43.45\text{ W/m}^2$ , the regulation target of the average temperature of supply and return water in secondary loop is  $45.2\text{ }^{\circ}\text{C}$ , and the control target of the primary loop return water temperature is  $49.1\text{ }^{\circ}\text{C}$ , and the control target of the secondary loop supply temperature is  $51\text{ }^{\circ}\text{C}$ . Different regulators have different tracking targets, and the heating substation is mainly regulated according to the secondary loop supply temperature as the target.

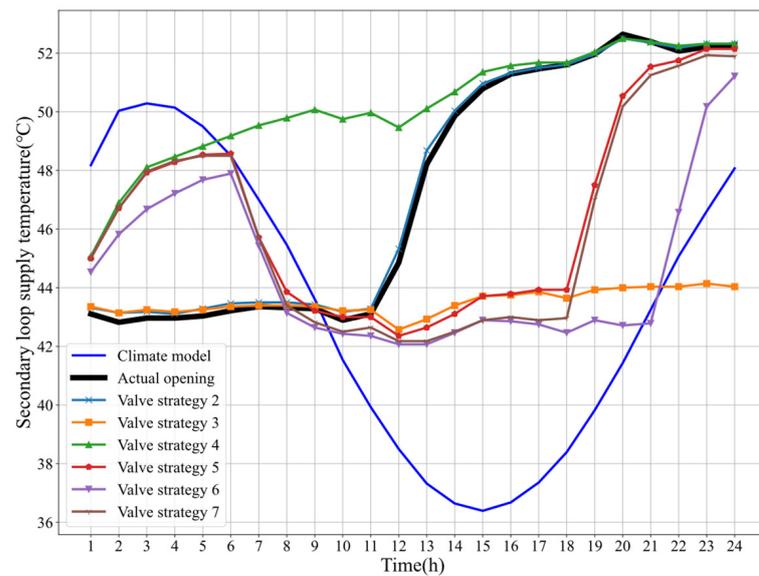
Table 4 lists six different regulation strategies. Valve strategy #1 is the control plan for the actual operation of the regulators of the heating substation on 4 January. Valve strategy #2 is to keep the valve opening unchanged at 15% the day before. Valve strategy #3 is to directly increase the opening to 20%. Valve strategies #4, #5, and #6 are all trying to follow changes with the outside temperature, which means regulating the valve with a larger opening when in the morning and evening, and regulating the valve to a small opening during the daytime. The according responses are presented in Figure 8.

**Table 4.** Six kinds of valve regulation strategies for the next 24 h.

Valve Regulation Strategy for the Next 24 h (%)												
	1	2	3	4	5	6	7	8	9	10	11	12
#1	15	15	14.8	15	15	15	15	15	15	15	15	18
#2	15	15	15	15	15	15	15	15	15	15	15	15
#3	20	20	20	20	20	20	20	20	20	20	20	20
#4	20	20	20	20	20	20	15	15	15	15	15	15
#5	18	18	18	18	18	18	10	10	10	10	10	10
#6	20	20	20	20	20	20	8	8	8	2	2	2
	13	14	15	16	17	18	19	20	21	22	23	24
#1	20	20	20	20	20	20	20	20	20	20	19.9	20
#2	15	15	15	15	15	15	15	15	15	15	15	15
#3	20	20	20	20	20	20	20	20	20	20	20	20
#4	15	15	15	15	15	15	18	18	18	18	18	18
#5	10	10	10	10	10	10	10	10	10	18	18	18
#6	2	2	2	2	2	2	18	18	18	18	18	18

Through the above temperature response model, operators can check the valve control strategy's actual temperature response before applying it.

The following are the three strategies, corresponding to valve strategy #1, valve strategy #2, and valve strategy #3, respectively, observing different valve control plans. First of all, it can be seen that the temperature prediction curve for valve strategy #1 is the same as the temperature response curve under the actual control plan. In the 24-h comparison of the predicted temperature response and the actual temperature response, the maximum prediction error of the model is  $0.47\text{ }^{\circ}\text{C}$ , the minimum prediction error is  $0.01\text{ }^{\circ}\text{C}$ , and the average prediction error is  $0.17\text{ }^{\circ}\text{C}$ , which verified the accuracy and reliability of the temperature response model. Secondly, valve strategy #2 is to maintain the valve opening 15% of the previous day unchanged, and the secondary loop supply temperature in the next 24 h is maintained at  $43\text{--}44\text{ }^{\circ}\text{C}$ . Finally, valve strategy #3 is to directly increase the valve opening from 15% to 20% and maintain this opening. It can be concluded that the corresponding temperature response curve gradually rises from around  $45\text{ }^{\circ}\text{C}$  and finally stabilizes at around  $52\text{ }^{\circ}\text{C}$ .



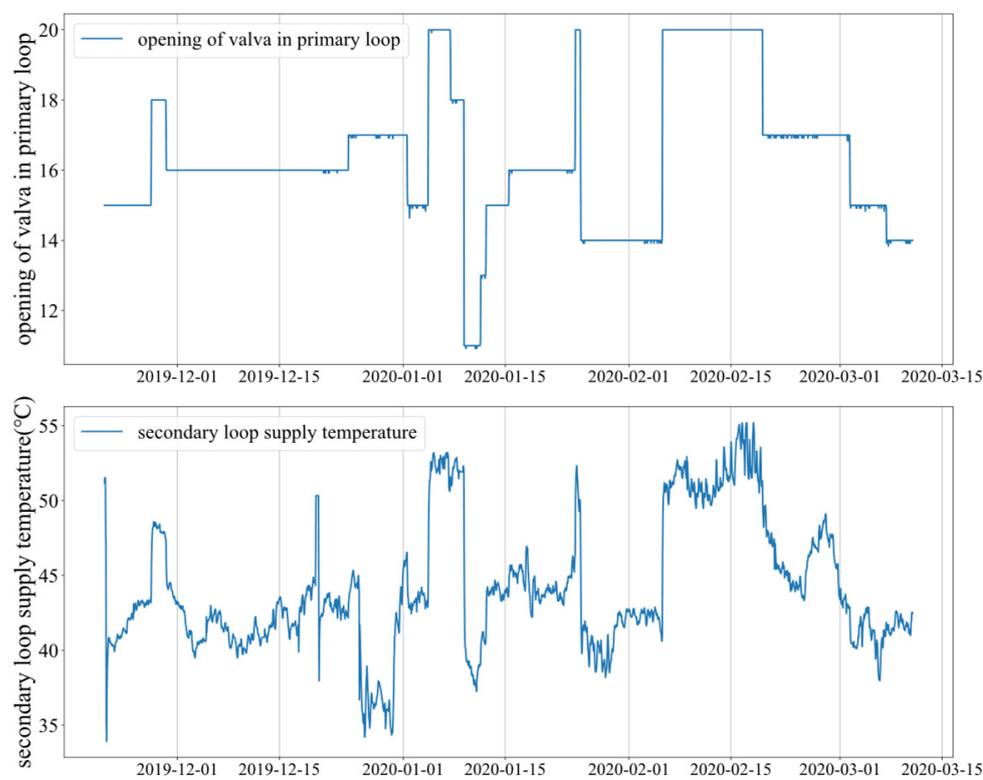
**Figure 8.** Temperature response comparison under different valve regulation strategies.

Here are the other three strategies, corresponding to valve strategy #4, valve strategy #5, and valve strategy #6, respectively, following the changes in the outside temperature. First of all, valve strategy #4 is to increase the valve opening of the heating substation to 20% at the beginning and then reduce it to 15% during the day, finally increasing the valve opening to 18% at night. The temperature response curve also correspondingly increases to around 48 °C, then gradually decreases to around 44 °C, and finally increases to 50 °C at night. Secondly, valve strategy #5 increases the control range based on regulation strategy 4, which reduces the valve opening to 10% during the day, and delays the increase in the valve by three hours at night; correspondingly, the temperature response of the secondary loop supply water increases from 45 °C to around 48 °C, then gradually decreases to 42 °C after the valve opening is reduced, and then gradually increases to 50 °C at night. Finally, valve strategy #6 tests the temperature response under extreme control actions. The valve opening was adjusted to 20% and 18% in the morning and evening, and the valve opening is reduced to 8% and 2% during the day, but Figure 8 shows that the corresponding secondary temperature response still maintains above 42 °C during the day, which is unreasonable according to expert knowledge and practical operation experience, so the temperature response prediction model may be invalid under such extreme condition.

The secondary loop supply temperature and the opening of valve changes in the experimental heating substation during the heating season of 2019–2020 are described in Figure 9 to analyze the failure of the temperature prediction response model with extreme conditions. Most of the time, the secondary loop supply temperature is above 40 °C, except from 28 December to 31 December when the secondary loop supply temperature is around 36 °C. In addition, as shown in Figure 9, it can be found that the valve opening of heating substation during 2019–2020 heating season ranges from 11% to 20%, and the valve opening range is basically maintained at 14% to 20%. The principle of machine learning is learning the internal relationships and laws from existing historical samples. However, for the training samples, the valve opening and the secondary loop supply temperature are rarely less than 14% and 40 °C, and there is even no condition where the valve opening is less than 10%. Considering the working conditions of the previous day and the future weather forecast, the reason for the failure in model response valve strategy #6 is that the high-dimensional sample space contained in the model has been exceeded during training.

To ensure the prediction accuracy and reliability of the temperature response prediction model, the valve regulation command, which is the model's input, should be kept in the working condition space of the training sample. For regulators, the temperature response

prediction model can predict the response temperature with different regulation strategies, which guide regulation and avoid the empirical regulation of “regulation-stable-re-regulation”.



**Figure 9.** Secondary loop supply temperature and valve opening in the 2019–2020 heating season.

#### 4. Conclusions

This research aims to solve the regulation problems in current heating substations. Firstly, a forecast model was established to calculate the supply temperature of the secondary loop in district heating systems using available information. Then, the relationship between the secondary loop’s supply temperature and the valve’s opening in the primary loop was investigated. Finally, a complete control chain of “collected information-predicted-regulate” is established. The results are discussed as following:

1. The prediction performance of the machine learning model is compared under different input step sizes and prediction step sizes. The XGBoost model with 72 steps of input and 24 steps of prediction is used as the temperature response prediction model, the average prediction error of which is 0.26%, which has a high prediction accuracy;
2. The XGBoost model with 72 steps of input and 24 steps of prediction is used to compare different valve opening control strategies. Based on the model, the valve control strategy of the heating substation is determined, which can realize the predictive control of the heating substation, improve the control accuracy of the heating substation, and reduce the dependence on manual experience;
3. The work of this paper was practically applied in a real district heating system in Zhengzhou, China. The final validation results showed that the adoption of the proposed regulation strategy resulted in a 5% improvement in system energy efficiency.

In this paper, only one historical operating condition of a single heating substation in one heating season is used for training, and the training sample space is limited, so the prediction performance of the model with extreme operating conditions outside the sample space is limited. It is necessary to explore further how to obtain a more widely distributed sample space in the heating substation.

**Author Contributions:** Conceptualization, Z.L. (Zhongbo Li) and X.L.; methodology, Z.L. (Zhongbo Li) and W.H.; software, Z.L. (Zhongbo Li) and W.H.; validation, Z.L. (Zhongbo Li) and W.H.; formal analysis, Z.L. (Zhongbo Li) and Z.L. (Zheng Luo); investigation, Z.L. (Zhongbo Li) and Z.L. (Zheng Luo); resources, Z.L. (Zhongbo Li) and N.Z.; data curation, Z.L. (Zhongbo Li) and N.Z.; writing—original draft preparation, Z.L. (Zheng Luo) and N.Z.; writing—review and editing, Z.L. (Zheng Luo) and N.Z.; visualization, Z.L. (Zheng Luo) and N.Z.; supervision, Z.L. (Zheng Luo) and N.Z.; project administration, X.L., E.F. and W.Z.; funding acquisition, X.L. and W.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Key R&D Program of China, grant number 2017YFA0700302 and this work was partly funded the by National Natural Science Foundation of China, grant number 51806190.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author. The data are not publicly available due to institutional confidentiality agreement.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Zhong, W.; Huang, W.; Lin, X.; Li, Z.; Zhou, Y. Research on data-driven identification and prediction of heat response time of urban centralized heating system. *Energy* **2020**, *212*, 118742. [[CrossRef](#)]
- The Building Energy Conservation Research Center of Tsinghua University. *Annual Research Report on the Development of Building Energy Efficiency in China 2019*; China Architecture and Building Press: Beijing, China, 2019.
- 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*, 1126–1139. [[CrossRef](#)]
- Wang, H.; Wang, H.; Zhu, T. A new hydraulic regulation method on district heating system with distributed variable-speed pumps. *Energy Convers. Manag.* **2017**, *147*, 174–189. [[CrossRef](#)]
- Stevanovic, V.D.; Zivkovic, B.; Prica, S.; Maslovaric, B.; Karamarkovic, V.; Trkulja, V. Prediction of thermal transients in district heating systems. *Energy Convers. Manag.* **2009**, *50*, 2167–2173. [[CrossRef](#)]
- Gu, J.; Wang, J.; Qi, C.; Yu, X.; Sundén, B. Analysis of a hybrid control scheme in the district heating system with distributed variable speed pumps. *Sustain. Cities Soc.* **2019**, *48*, 101591. [[CrossRef](#)]
- Gustafsson, J.; Delsing, J.; van Deventer, J. Experimental evaluation of radiator control based on primary supply temperature for district heating substations. *Appl. Energy* **2011**, *88*, 4945–4951. [[CrossRef](#)]
- Henze, G.P.; Floss, A.G. Evaluation of temperature degradation in hydraulic flow networks. *Energy Build.* **2011**, *43*, 1820–1828. [[CrossRef](#)]
- Mendes, N.; Oliveira, G.; Araújo, H. Building thermal performance analysis by using matlab/simulink. In Proceedings of the Seventh International IBPSA Conference, Rio de Janeiro, Brazil, 13–15 August 2001.
- Karlsson, J.; Wadsö, L.; Öberg, M. A conceptual model that simulates the influence of thermal inertia in building structures. *Energy Build.* **2013**, *60*, 146–151. [[CrossRef](#)]
- Yu, Z.; Haghighat, F.; Fung, B.C.M.; Yoshino, H. A decision tree method for building energy demand modeling. *Energy Build.* **2010**, *42*, 1637–1646. [[CrossRef](#)]
- Touzani, S.; Granderson, J.; Fernandes, S. Gradient boosting machine for modeling the energy consumption of commercial buildings. *Energy Build.* **2018**, *158*, 1533–1543. [[CrossRef](#)]
- Magnier, L.; Haghighat, F. Multiobjective optimization of building design using TRNSYS simulations, genetic algorithm, and Artificial Neural Network. *Build. Environ.* **2010**, *45*, 739–746. [[CrossRef](#)]
- Machado, J.E.; Cucuzzella, M.; Scherpen, J.M.A. Modeling and passivity properties of multi-producer district heating systems. *Automatica* **2022**, *142*, 110397. [[CrossRef](#)]
- Xu, B.; Huang, A.; Fu, L.; Di, H. Simulation and analysis on control effectiveness of TRVs in district heating systems. *Energy Build.* **2011**, *43*, 1169–1174. [[CrossRef](#)]
- Chicherin, S.; Anvari-Moghaddam, A. Adjusting heat demands using the operational data of district heating systems. *Energy* **2021**, *235*, 121368. [[CrossRef](#)]
- Wang, R.; Wang, Q.; Wei, C.; Cong, M.; Zhou, Z.; Ni, L.; Liu, J.; Wei, W. A Thermo-Hydraulic couplings model for residential heating system based on Demand-side Regulation: Development and calibration. *Energy Build.* **2022**, *256*, 111667. [[CrossRef](#)]
- Zheng, J.; Zhou, Z.; Zhao, J.; Hu, S.; Wang, J. Effects of intermittent heating on an integrated heat and power dispatch system for wind power integration and corresponding operation regulation. *Appl. Energy* **2021**, *287*, 116536. [[CrossRef](#)]
- Zhao, J.; Lyu, L.; Han, X. Operation regulation analysis of solar heating system with seasonal water pool heat storage. *Sustain. Cities Soc.* **2019**, *47*, 101455. [[CrossRef](#)]

20. Sun, C.; Liu, Y.; Gao, X.; Wang, J.; Yang, L.; Qi, C. Research on control strategy integrated with characteristics of user's energy-saving behavior of district heating system. *Energy* **2022**, *245*, 123214. [[CrossRef](#)]
21. Wang, D.; Zhi, Y.-Q.; Jia, H.-J.; Hou, K.; Zhang, S.-X.; Du, W.; Wang, X.-D.; Fan, M.-H. Optimal scheduling strategy of district integrated heat and power system with wind power and multiple energy stations considering thermal inertia of buildings under different heating regulation modes. *Appl. Energy* **2019**, *240*, 341–358. [[CrossRef](#)]
22. Cadau, N.; Lorenzi, A.D.; Gambarotta, A.; Morini, M.; Saletti, C. A Model-in-the-Loop application of a Predictive Controller to a District Heating system. *Energy Procedia* **2018**, *148*, 352–359. [[CrossRef](#)]
23. Bojic, M.; Trifunovic, N. Linear programming optimization of heat distribution in a district-heating system by valve adjustments and substation retrofit. *Build. Environ.* **2000**, *35*, 151–159. [[CrossRef](#)]
24. Turski, M.; Sekret, R. Buildings and a district heating network as thermal energy storages in the district heating system. *Energy Build.* **2018**, *179*, 49–56. [[CrossRef](#)]
25. Ruseljuk, P.; Dedov, A.; Hlebnikov, A.; Lepiksaar, K.; Volkova, A. Comparison of District Heating Supply Options for Different CHP Configurations. *Energies* **2023**, *16*, 603. [[CrossRef](#)]
26. García-Céspedes, J.; Herms, I.; Arnó, G.; de Felipe, J.J. Fifth-Generation District Heating and Cooling Networks Based on Shallow Geothermal Energy: A review and Possible Solutions for Mediterranean Europe. *Energies* **2023**, *16*, 147. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.