energies

MDPI

Article

Systematic Method for the Energy-Saving Potential Calculation of Air Conditioning Systems via Data Mining. Part II: A Detailed Case Study

Rongjiang Ma [1,2], **Shen Yang** [1,3,*], **Xianlin Wang** [1,4,5], **Xi-Cheng Wang** [4,5], **Ming Shan** [1], **Nanyang Yu** [2] and **Xudong Yang** [1]

1　Department of Building Science, Tsinghua University, Beijing 100084, China; rongjiangma@139.com (R.M.); wangxianlingree@126.com (X.W.); mshan@mail.tsinghua.edu.cn (M.S.); xyang@tsinghua.edu.cn (X.Y.)
2　School of Mechanical Engineering, Southwest Jiaotong University, Chengdu 610031, China; rhinos@126.com
3　Human-Oriented Built Environment Lab, School of Architecture, Civil and Environmental Engineering, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland
4　State Key Laboratory of Air-Conditioning Equipment and System Energy Conservation, Zhuhai 519070, China; xichengwanggree@163.com
5　Gree Electric Appliances, Inc. of Zhuhai, Zhuhai 519070, China
*　Correspondence: yangshen13@tsinghua.org.cn

Abstract: Increased data monitoring enables the energy-efficient operation of air-conditioning systems via data-mining. The latter is projected to have lesser consumption but more comprehensive diagnosis than traditional methods. Following the companion paper that proposed a systematic method for energy-saving potential calculations via data-mining, this article presents a detailed case study in an ice-storage air-conditioning system by employing the proposed method. Raw data were preprocessed prior to recognizing the constant- and variable-speed devices in the system. Classification and regression tree algorithms were utilized to identify the operating modes of the system. The regression models between the energy-consumption and operating-state parameters of the nine pumps and two chillers were fitted. Furthermore, the constraints pertaining to system operation were summarized. From the results, the particle swarm optimization method was applied to elucidate the benchmark energy cost and the consequent cost savings potential. The cost savings potential for the chiller plant room during the investigation duration of 59 d reached as high as 24.03%. The case study demonstrates the feasibility, effectiveness, and stability of the systematic approach. Further studies can facilitate the development of corresponding control strategies based on the potential analysis results, to investigate better optimization algorithm, and visualize the analysis process.

Keywords: energy-saving potential; data-mining; recognition; optimization; operational data

1. Introduction

As reviewed in a companion paper [1], data mining can be a superior approach for diagnosing the energy-saving potential of air-conditioning systems, particularly in the era of data explosion. Compared to the traditional observation/question, test/calculation, and identification/resolution (OTI) method, energy-saving diagnosis based on data mining exhibits several advantages, including a more comprehensive analysis of energy-saving diagnosis, reduced intervention of professional researchers, and less time. The companion paper [1] proposed a systematic energy-saving diagnosis method for air conditioning systems via data-mining. The proposed method consisted of seven steps: (1) data collection, (2) data preprocessing, (3) recognition of variable frequency equipment, (4) recognition of system operation mode, (5) regression analysis of energy-consumption data, (6) constraint analysis of the system during operation, and (7) analysis of energy-saving potential. To validate the proposed method and test its applicability and feasibility for application in complicated air conditioning systems, this study mainly focuses on the technical details

of the method and the application of the method in a specific air conditioning system for energy-saving diagnosis.

Previous studies have investigated the applicability of data-mining technologies in energy-consumption-related investigations into air-conditioning systems and building energy systems [2]. For instance, an artificial neural network (ANN) [3] and support vector machine (SVM) [4] have been employed to diagnose chillers. Moreover, Gaussian process regression [5] and classification and regression trees (CART) [6] have been utilized to diagnose air handling units (AHUs). ANN [7] and an algorithm based on the recursive deterministic perceptron neural network [8] have been utilized to detect pumps; symbolic aggregate approximation (SAX) have been applied to identify the operation patterns of chillers [9] and heating, ventilating and air conditioning (HVAC) systems [10]. Few studies have focused on the application of various data mining technologies, such as response surface methodology (RSM) and neural network (NN), in the diagnosis and optimization of specific components of HVAC systems and building energy systems. Li et al. [11] applied clustering analysis and association mining to identify energy consumption patterns of a variable refrigerant flow system. Neural network (NN) showed capability in optimizing specific components, such as fluid in the refrigeration system [12]. Response surface methodology (RSM) was applied for the optimization of specific building energy systems, such as solar collector and reactor [13,14]. Global data mining, together with geographic information, has enabled country-wide optimization of building insulation for energy saving and mitigation of emissions [15]. However, the aforementioned case studies only pertain to a specific step or optimization algorithm included in our systematic method. Newly proposed methods for energy saving analysis of HVAC systems, particularly the ones focusing on systematic optimization, should be validated using case studies. Previous studies that proposed new optimization methods for HVAC systems [16], building energy systems [17], or specific components [18] were normally accompanied by case studies for validation. Therefore, a comprehensive and detailed case study employing the new systematic data-mining-based methodology is necessary to illustrate the method's feasibility.

This paper presents the second of two publications proposing a systematic methodology to elucidate the energy-saving potential of an air conditioning system based on data-mining. Following the proposed steps in the companion paper [1], a detailed case study was carried out in an air conditioning system coupled with an ice storage system with an air conditioning area of 30,000 m^2. The details of specific data-mining technologies in each step were introduced, and the energy-saving potential was calculated and analyzed using the systematic method.

2. Methodology

2.1. The Studied System

A five-floor commercial building was equipped with an air-conditioning system by employing the new method. The building was located in Shenzhen, China, with a 30,000 m^2 air conditioning area. The system was coupled with an ice-storage system (circulation medium: glycol-water solution) to exploit the lower electricity price at night (Table A1 in Appendix A). As shown in Figure 1, the coupled system consisted of two chillers (Ch-1 and Ch-2), three chilled-water pumps (ChWP-1, ChWP-2, and ChWP-3), three condensing-water pumps (CWP-1, CWP-2, and CWP-3), three glycol water pumps (GWP-1, GWP-2, and GWP-3), twelve ice-storage tanks, two plate heat exchangers, six cooling towers, and five air-handling units.
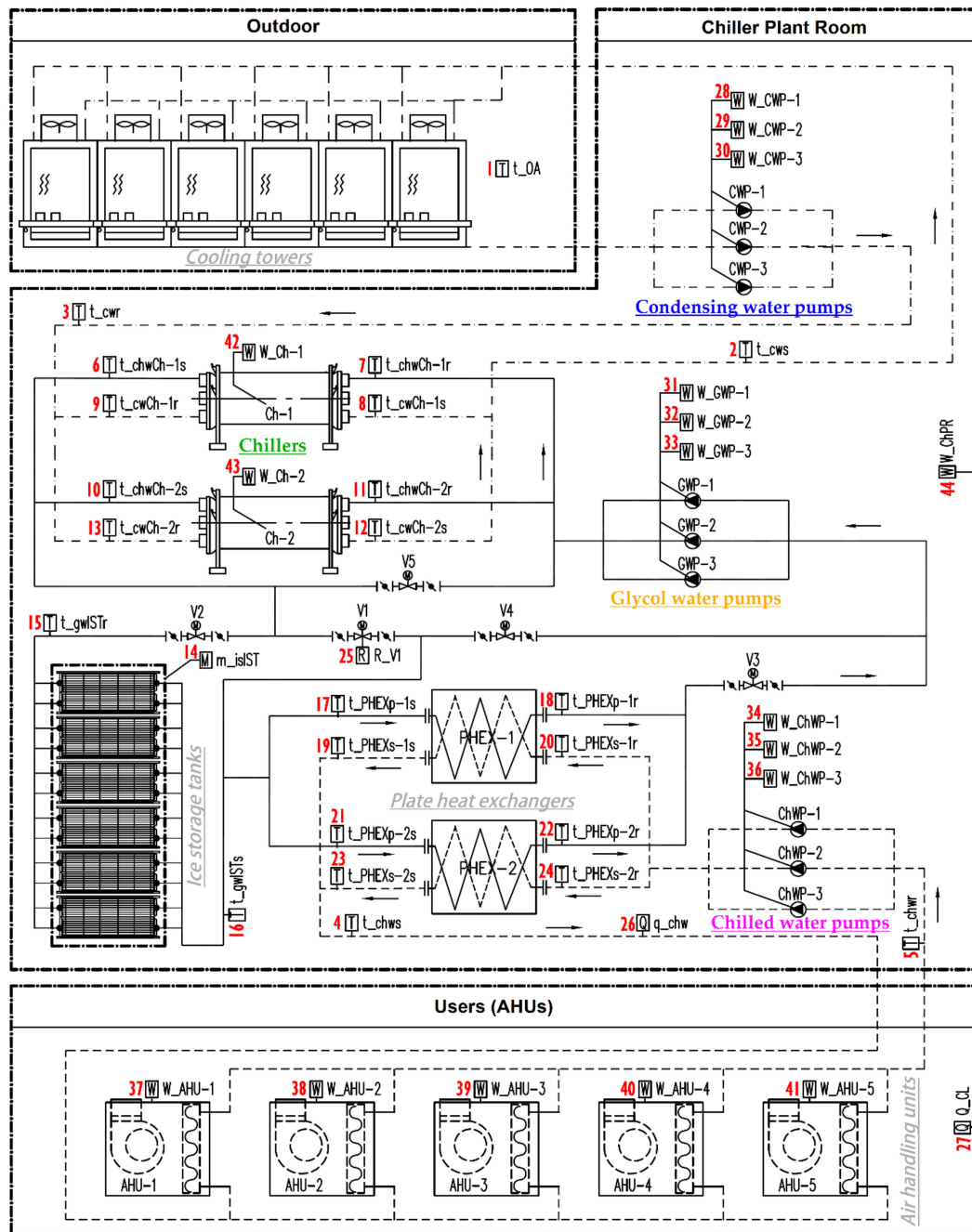
Figure 1. Scheme of the studied air-conditioning system with the monitoring locations of the operation status.

The system had four operating modes, excluding the shutdown mode. Each mode was controlled based on the logic detailed in Table 1 for on/off control or the modulation of the chiller(s) and five valves (V1 to V5).

Table 1. System operating modes and corresponding control logic.

Operating Modes	Codes	Device Status					
		Chiller(s)	V1	V2	V3	V4	V5
Ice build	M1	On	Off	On	Off	On	Off
Cooling by chiller(s) only	M2	On	On	Off	On	Off	Off
Cooling by ice only	M3	Off	Off	On	On	Off	On
Cooling by ice with chiller(s)	M4	On	Modulate	On	On	Modulate	Off

2.2. Data Collection

A comprehensive data-monitoring platform was established for the air-conditioning system to monitor the electricity consumption of each component and the operating status parameters of the system. The available monitoring points are indicated in Figure 1, and all the points are summarized in Table 2, along with detailed information.

Table 2. Parameters monitored in the case system.

No. in Figure 1	Identifiers in Figure 1	Parameter Meaning and Unit	Symbol
1	t_OA	Outdoor air temperature (°C)	t_{OA}
2	t_cws	Temperature of supply condensing water in main pipe (°C)	t_{cws}
3	t_cwr	Temperature of return condensing water in main pipe (°C)	t_{cwr}
4	t_chws	Temperature of supply chilled water in main pipe (°C)	t_{chws}
5	t_chwr	Temperature of return chilled water in main pipe (°C)	t_{chwr}
6	t_chwCh-1 s	Temperature of supply chilled water in chiller #1 (°C)	$t_{chwCh\text{-}1\,s}$
7	t_chwCh-1r	Temperature of return chilled water in chiller #1 (°C)	$t_{chwCh\text{-}1r}$
8	t_cwCh-1 s	Temperature of supply condensing water in chiller #1 (°C)	$t_{cwCh\text{-}1\,s}$
9	t_cwCh-1r	Temperature of return condensing water in chiller #1 (°C)	$t_{cwCh\text{-}1r}$
10	t_chwCh-2 s	Temperature of supply chilled water in chiller #2 (°C)	$t_{chwCh\text{-}2\,s}$
11	t_chwCh-2r	Temperature of return chilled water in chiller #2 (°C)	$t_{chwCh\text{-}2r}$
12	t_cwCh-2 s	Temperature of supply condensing water in chiller #2 (°C)	$t_{cwCh\text{-}2\,s}$
13	t_cwCh-2r	Temperature of return condensing water in chiller #2 (°C)	$t_{cwCh\text{-}2r}$
14	m_isIST	Amount of ice in inventory (metric tons)	m_{isIST}
15	t_gwISTr	Temperature of return glycol water to ice storage tanks (°C)	t_{gwISTr}
16	t_gwISTs	Temperature of supply glycol water from ice storage tanks (°C)	t_{gwISTs}
17	t_PHEXp-1 s	Temperature of supply water to premier side of plate heat exchanger #1 (°C)	$t_{PHEXp\text{-}1\,s}$
18	t_PHEXp-1r	Temperature of return water from premier side of plate heat exchanger #1 (°C)	$t_{PHEXp\text{-}1r}$
19	t_PHEXs-1 s	Temperature of supply water from secondary side of plate heat exchanger #1 (°C)	$t_{PHEXs\text{-}1\,s}$
20	t_PHEXs-1r	Temperature of return water to secondary side of plate heat exchanger #1 (°C)	$t_{PHEXs\text{-}1r}$
21	t_PHEXp-2 s	Temperature of supply water to premier side of plate heat exchanger #2 (°C)	$t_{PHEXp\text{-}2\,s}$
22	t_PHEXp-2r	Temperature of return water from premier side of plate heat exchanger #2 (°C)	$t_{PHEXp\text{-}2r}$
23	t_PHEXs-2 s	Temperature of supply water from secondary side of plate heat exchanger #2 (°C)	$t_{PHEXs\text{-}2\,s}$
24	t_PHEXs-2r	Temperature of return water to secondary side of plate heat exchanger #2 (°C)	$t_{PHEXs\text{-}2r}$
25	R_V1	Opening of valve #1 (%)	R_{V1}
26	q_chw	Chilled water flow in main pipe (L/s)	q_{chw}
27	Q_CL	Cooling load of the system (USRT)	Q_{CL}
28	W_CWP-1	Energy-consumption of condensing water pump #1 (kWh)	$W_{CWP\text{-}1}$
29	W_CWP-2	Energy-consumption of condensing water pump #2 (kWh)	$W_{CWP\text{-}2}$
30	W_CWP-3	Energy-consumption of condensing water pump #3 (kWh)	$W_{CWP\text{-}3}$
31	W_GWP-1	Energy-consumption of glycol water pump #1 (kWh)	$W_{GWP\text{-}1}$
32	W_GWP-2	Energy-consumption of glycol water pump #2 (kWh)	$W_{GWP\text{-}2}$
33	W_GWP-3	Energy-consumption of glycol water pump #3 (kWh)	$W_{GWP\text{-}3}$
34	W_ChWP-1	Energy-consumption of chilled water pump #1 (kWh)	$W_{ChWP\text{-}1}$
35	W_ChWP-2	Energy-consumption of chilled water pump #2 (kWh)	$W_{ChWP\text{-}2}$
36	W_ChWP-3	Energy-consumption of chilled water pump #3 (kWh)	$W_{ChWP\text{-}3}$
37	W_AHU-1	Energy-consumption of air handling unit #1 (kWh)	$W_{AHU\text{-}1}$
38	W_AHU-2	Energy-consumption of air handling unit #2 (kWh)	$W_{AHU\text{-}2}$
39	W_AHU-3	Energy-consumption of air handling unit #3 (kWh)	$W_{AHU\text{-}3}$
40	W_AHU-4	Energy-consumption of air handling unit #4 (kWh)	$W_{AHU\text{-}4}$
41	W_AHU-5	Energy-consumption of air handling unit #5 (kWh)	$W_{AHU\text{-}5}$
42	W_Ch-1	Energy-consumption of chiller #1 (kWh)	$W_{Ch\text{-}1}$
43	W_Ch-2	Energy-consumption of chiller #2 (kWh)	$W_{Ch\text{-}2}$
44	W_ChPR	Energy-consumption of chiller plant room (kWh)	W_{ChPR}

This study collected data from 22 July 2011 to 20 August 2013 at intervals of 1 h. All the data were utilized in this study for preprocessing and subsequent analyses. To better explain the specific types of data monitored by this study, Table A2 of Appendix A provides the raw data for a full day (18 August 2013).

2.3. Data Preprocessing

The raw data were preprocessed to meet the needs of the subsequent analysis following the steps detailed in the subsequent sections.

2.3.1. Missing Data Preprocessing

- Listwise deletion: We deleted the data obtained before 22 June 2013, owing to abundant missing values in the data from that period. This was done to ensure the continuity, isometry, and completeness of the data. Consequently, 1416 pairs of continuous time-series data for 59 consecutive days from 22 June to 19 August 2013 were retained for the recognition of system operation mode in Section 2.5 and the energy cost-savings potential analysis in Section 2.8.
- Pairwise deletion: In the regression analysis (Section 2.6) of energy consumption and flow rate for pumps, we deleted the energy-consumption data and the paired flow rate data in cases where either one or both values were missing. The same approach was applied to the data used in the regression analysis for the chillers.

2.3.2. Data Cleaning

- Duplicate data cleaning: Consider the data obtained at 17:59:00 on 18 August 2013 (see Table A2). Because the data were recorded at 1 h intervals, the data obtained at 17:59:00 was considered to be a duplicate of the approaching hour (18:00:00) and was deleted.
- Data cleaning during the shutdown state: To faithfully reflect the distribution of energy consumption in the operating state, before calculating the numerical characteristics of the energy-consumption data of the devices in Section 2.4, the data obtained during the shutdown state (i.e., when the energy consumption was zero) were deleted. The same approach was applied to the data used in Section 2.6 for regression analysis.

2.3.3. Data Extending

- Appending the code of system operation mode to the dataset: Following the recognition of the system operation mode (see Section 2.5), we added a column into the dataset with the corresponding operation-mode code (i.e., M1 to M4 in Table 1 and M0 for the shutdown mode) to facilitate the filtering of data by operation mode in the relevant analysis. For example, the chillers were regressed by cooling (corresponding to M2 and M4) and ice building (corresponding to M1) modes in Section 2.6.
- Temperature difference data extension: Typically, temperature differences are not directly captured and recorded during the data collection period but are frequently used in the analysis. Therefore, it is necessary to add temperature difference data to facilitate direct recall for the relevant analysis (e.g., regression analysis in Section 2.7). For instance, we computed the temperature difference between the supply and the return chilled water ($\Delta t_{\mathrm{chw}} = t_{\mathrm{chwr}} - t_{\mathrm{chws}}$) as a new variable ($\Delta t_{\mathrm{chw}}$) and extended it to the case dataset. Similarly, for the temperature difference of the water supply and return chilled water of the chiller ($\Delta t_{\mathrm{chwCh}}$), the temperature difference between the water supply and return condensing water of the chiller (Δt_{cwCh}) were calculated and extended to the dataset.

2.3.4. Data Transformation

- Transformation of the units of measurement: The unit of the cooling load of the system in the case data was the US refrigeration ton (USRT). This unit pertains to the cooling load (power nature) rather than the cooling quantity demand (energy

nature); the former cannot be used directly in the subsequent analysis and calculation. Hence, the unit of USRT was converted to the système international (SI) unit of kilowatts (kW). Because the data collection interval was 1 h, each value corresponded to the cooling quantity demand for that period. Therefore, the unit was further converted to kilowatt-hour (kWh). In the case of chilled water flow (q_{chw}), where there were no more variables of the same type, the unit of liters per second (L/s) was converted to cubic meters per hour (m^3/h) to avoid introducing too many conversion factors in the subsequent analysis and calculation.

- Time interval labeling: For easier identification and more efficient data processing, the time interval was marked as period *i*. The one-to-one correspondences between them are shown in Table A1 of Appendix A.

2.4. Recognition of Variable Speed Equipment

The regression analysis in Section 2.6 requires the matching of different fitting formulas for variable- and constant-speed motor-driven equipment; however, the variable- or constant-speed nature of equipment was not directly available in the monitoring data. Therefore, we applied the coefficient of the median, defined as the ratio of the difference between the maximum and median to the range, proposed in our previous study [19]. The coefficient of the median aids in recognizing the two-speed types of motor-driven equipment using their energy-consumption data. It was evident from the recognition results that the four pumps (ChWP-1, ChWP-2, GWP-1, and GWP-2) and all the AHUs were variable-speed equipment, whereas the remaining five pumps (CWP-1, CWP-2, CWP-3, ChWP-3, and GWP-3) and the two chillers (Ch-1 and Ch-2) were constant-speed devices.

2.5. Recognition of System-Operation Mode

The operation mode was not recorded for each period in the dataset. Therefore, it is necessary to recognize or classify the system operation modes before relevant analysis.

As previously stated, in the studied air-conditioning system, there were five distinct operation-modes: (a) shutdown (operation-mode code: M0), (b) ice build (M1), (c) cooling by chiller(s) only (M2), (d) cooling by ice only (M3), and (e) cooling by both chiller(s) and ice (M4). Moreover, the different operation modes of the system follow the clear control logic shown in Table 1. While the data from the chillers and valve V1 can recognize modes M1 to M4, they are insufficient to recognize mode M0, which is easily confused with M3. Therefore, the energy consumption of another energy-consuming equipment is incorporated to ensure the correct recognition of mode M0. The energy-consumption types of the involved devices, corresponding to different operation modes, are shown in Table 3. The use of electricity at night versus peak daytime hours can lead to large savings in energy bills. This proves that the system primarily produces ice during the nighttime low-tariff hours (23:00 to 07:00) when the building is closed to the public, and no cooling is required. The period *i* is therefore added to the mode recognition, which may be effective in recognizing mode M1.

Table 3. Energy-consumption types of devices for different operation modes.

Operation-Mode Codes	Chiller (s)	Air Handling Unit (s)	Chilled Water Pump (s)	Glycol Water Pump (s)	Condensing Water Pump (s)
M0	0 [1]	0	0	0	0
M1	>0 [2]	0	0	>0	>0
M2	>0	>0	>0	>0	>0
M3	0	>0	>0	>0	0
M4	>0	>0	>0	>0	>0

[1] "0" indicates that the corresponding device is in a shutdown state. [2] ">0" indicates that the device is running.

Classification is a method of building a categorization model by summarizing and refining the patterns contained in existing data. Moreover, decision tree induction is a classical classification method with high accuracy and efficiency. To better elucidate the data and incorporate background knowledge about it when building a decision tree, the classification and regression trees (CART) algorithm [6] with good interactivity was chosen to build the decision tree model for the study. CART uses a greedy method in which the decision tree is constructed using a top-down recursive partitioning approach. For the classification of numerical input variables, CART measures the heterogeneity of the output variables by calculating the Gini index (*Gini*) [20] for each input variable and selects the splitting variable that maximizes the reduction in heterogeneity ($\Delta Gini$). This variable and its split-point together form the splitting criterion. This procedure is repeated until the splitting criteria for recognizing all operation modes are acquired, at the end of which the decision tree construction is complete.

2.6. Regression Analysis of Energy-Consumption Data

The purpose of the regression analysis is to quantify the relationship between the energy consumption of each energy-consuming device and its operation-state parameters. The obtained fitting model serves to reduce energy consumption by optimizing the operation-state parameters while meeting the system load demands. The energy-consumption data of the cooling towers and operation-state parameters for AHUs were not available in the dataset; therefore, follow-up analysis was performed only for the energy-consuming equipment in the chiller plant room, which included the nine pumps and two chillers.

2.6.1. Fitting Models for Regression Analysis

The nine pumps in the case study are categorized into constant- and variable-speed pumps based on the classification in Section 2.4, corresponding to different fitting models. Based on fluid dynamics, the fitting model of the energy consumption and flow rate for the constant-speed centrifugal pumps is as follows:

$$W = \beta_0 + \beta_1 q, \tag{1}$$

where W represents the energy consumption of pump, kWh; q represents the flow rate of the pump, m^3/h; and β_0 and β_1 are the two fitting coefficients and are dimensionless.

The fitting model for variable-speed centrifugal pumps can be determined as follows:

$$W = \beta_1 (q + \beta_0)^3. \tag{2}$$

Several factors influence chiller-operating performance and energy consumption; an accurate theoretical model can be developed via an in-depth analysis of chilling principles. However, this study of chiller-operating performance aims to exploit actual operation data to fit a real or near-real model that can predict energy consumption in the following energy-saving potential calculations. In some related studies, the mathematical models from two well-known simulation software programs in the field, EnergyPlus (9.3.0, National Renewable Energy Laboratory, Golden, CO, USA) and TRNSYS (v. 17, Thermal Energy System Specialists, LLC, Madison, WI, USA), were used directly or with appropriate modifications based on the studies [21]. Herein, an improved model that can be effectively fitted as proposed in our previous study [22] is applied as follows:

$$
\begin{aligned}
W_{\text{Ch}} = \ & \beta_0 + \beta_1 Q_{\text{Ch}} + \beta_2 t_{\text{cwChr}} + \beta_3 \Delta t_{\text{chwCh}} + \beta_4 \Delta t_{\text{cwCh}} + \beta_5 (Q_{\text{Ch}})^2 + \\
& \beta_6 (t_{\text{cwChr}})^2 + \beta_7 (\Delta t_{\text{chwCh}})^2 + \beta_8 (\Delta t_{\text{cwCh}})^2 + \beta_9 Q_{\text{Ch}} t_{\text{cwChr}} + \\
& \beta_{10} Q_{\text{Ch}} \Delta t_{\text{chwCh}} + \beta_{11} Q_{\text{Ch}} \Delta t_{\text{cwCh}} + \beta_{12} t_{\text{cwChr}} \Delta t_{\text{chwCh}} + \beta_{13} t_{\text{cwChr}} \Delta t_{\text{cwCh}} + \\
& \beta_{14} \Delta t_{\text{chwCh}} \Delta t_{\text{cwCh}} + \beta_{15} t_{\text{chwChs}} + \beta_{16} (t_{\text{chwChs}})^2 + \beta_{17} Q_{\text{Ch}} t_{\text{chwChs}} + \\
& \beta_{18} t_{\text{cwChr}} t_{\text{chwChs}} + \beta_{19} \Delta t_{\text{chwCh}} t_{\text{chwChs}} + \beta_{20} \Delta t_{\text{cwCh}} t_{\text{chwChs}}
\end{aligned}
\tag{3}
$$

where W_{Ch} represents the energy consumption of the chiller, kWh; Q_{Ch} represents the cooling capacity of the chiller, kWh; t_{cwChr} represents the temperature of return condensing water of chiller, °C; Δt_{chwCh} represents temperature difference of water supply and the return chilled water of the chiller, °C; Δt_{cwCh} represents the temperature difference of water supply and the return condensing water of the chiller, °C; t_{chwChs} represents the temperature of water supply chilled water of the chiller, °C; β_0 to β_{20} represent the fitting coefficients and are dimensionless.

2.6.2. Data Preparation for Regression Analysis

Abnormal data have already been eliminated in the data preprocessing step (Section 2.3). However, the energy consumption data of each of the nine pumps were recorded in the case dataset. The corresponding flow rate was recorded only for the main pipe of the chilled water loop (q_{chw}), without a separate flow rate available for each pump. Therefore, it is crucial for the pump data to be processed in the following ways for regression analysis:

- Filter out the flow rate of the individual pump. Considering the chilled water flow rate (q_{chw}) as an example, the data are filtered out with only one chilled water pump in operation from all the data with flow rates (q_{chw}) exceeding zero. The filtered flow data and the corresponding pump energy consumption in the order of pump numbers were used for the respective regression analyses. For instance, a period when ChWP-1 runs while ChWP-2 and ChWP-3 are shut down is determined, chilled water flow rate (q_{chw}) is marked as the flow rate of ChWP-1 (q_{chw-1}), and the process is repeated to filter out the flow rate data for all ChWP-1. Subsequently, these filtered data are organized into a subset of ChWP-1 for regression analysis. The relevant data subsets for ChWP-2 and ChWP-3 can be collated separately following the same procedure.
- Calculate the flow rate of the glycol water (q_{gw}) and the flow rate of the condensing water (q_{cw}). In this study, ignoring the heat transfer loss of the system, $Q_{chw} = Q_{gw}$, the glycol water flow rate (q_{gw}) can be calculated using the chilled water flow rate (q_{chw}), the temperature difference between supply and return chilled water (Δt_{chw}), and the temperature difference between supply and return glycol water solution (Δt_{gw}) in the operating modes of M2, M3, and M4. Similarly, $Q_{cw} = Q_{Ch} + W_{Ch} = Q_{chw} + W_{Ch}$, the condensing water flow rate (q_{cw}) can be calculated using the chilled water flow rate (q_{chw}), the temperature difference between supply and return chilled water (Δt_{chw}), the energy consumption of the chillers, and the temperature difference between supply and return condensing water (Δt_{cw}) in the operating mode of M2. Finally, using the first method, the relevant data subsets for each glycol water pump and condensing water pump can be collated separately.

In addition, for chillers, after the preprocessing discussed in Section 2.3, the data for the regression analysis no longer require further processing.

2.7. Constraint Analysis of System during Operation

The case system is subjected to the following constraints during operation, which will also be taken into account in the energy-saving potential analysis.

2.7.1. Constraint on Supply and Demand of Cooling

During the cooling period in the operation modes of M2, M3, and M4, the cooling capacity supplied by the chiller plant room must be sufficient to meet the cooling load of the case system, described by the following equation:

$$Q_{CL}(i) = Q_{Ch}(i) + Q_{IST}^{cr}(i), \tag{4}$$

where $Q_{CL}(i)$ represents the cooling load of the system during the period i, kWh; $Q_{Ch}(i)$ represents the cooling capacity supplied by chillers during period i, kWh; and $Q_{IST}^{cr}(i)$ represents the cooling capacity supplied by ice in the ice storage tanks during period i, kWh.

2.7.2. Constraint on Cooling Capacity of Chillers

In the operation modes of M1, M2, and M4, the cooling capacity supplied by chillers should not exceed the maximum cooling capacity of chillers at any time.

$$0 \leq Q_{\text{Ch}}(i) \leq Q_{\text{Ch}}^{\max}(i), \tag{5}$$

where $Q_{\text{Ch}}^{\max}(i)$ represents the maximum cooling capacity of chiller during period i, kWh.

2.7.3. Constraints on ISTs

The sum of the accumulation of the cooling capacity in the ISTs and the current remaining cooling capacity should not exceed the maximum accumulation of cooling capacity of ISTs during any time:

$$0 \leq Q_{\text{IST}}^{\text{ac}}(i) + Q_{\text{IST}}^{\text{re}}(i) \leq Q_{\text{IST}}^{\max}, \tag{6}$$

where $Q_{\text{IST}}^{\text{ac}}(i)$ represents the accumulation of cooling capacity in the ISTs during period i, kWh; $Q_{\text{IST}}^{\text{re}}(i)$ represents the current remaining cooling capacity in the ISTs at the beginning of period i, kWh; $Q_{\text{IST}}^{\max}$ represents the maximum accumulation of cooling or cooling release of ISTs during one cooling storage and release cycle, kWh.

In the operation mode of M1, the accumulation of cooling capacity in the ISTs should be equal to the cooling capacity supplied by the chillers during any time, ignoring the heat transfer loss.

$$Q_{\text{IST}}^{\text{ac}}(i) = Q_{\text{Ch}}(i). \tag{7}$$

Moreover, in the operation mode of M1, ISTs are constrained by their operation efficiency; the accumulation of cooling capacity cannot exceed the maximum cooling storage speed in the current period.

$$0 \leq Q_{\text{IST}}^{\text{ac}}(i) \leq Q_{\text{IST}}^{\text{ac},\max}(i), \tag{8}$$

where $Q_{\text{IST}}^{\text{ac},\max}(i)$ represents the maximum cooling storage speed of ISTs during period i, kWh.

When cooling by ice (i.e., the operation mode of M3 and M4), the cooling release of ISTs at any time should not exceed the remaining cooling capacity during that time.

$$0 \leq Q_{\text{IST}}^{\text{cr}}(i) \leq Q_{\text{IST}}^{\text{re}}(i), \tag{9}$$

where $Q_{\text{IST}}^{\text{cr}}(i)$ represents the cooling release of ISTs during period i, kWh.

Similarly, the cooling release of ISTs cannot exceed the maximum cooling release speed in the current period.

$$0 \leq Q_{\text{IST}}^{\text{cr}}(i) \leq Q_{\text{IST}}^{\text{cr},\max}(i), \tag{10}$$

where $Q_{\text{IST}}^{\text{cr},\max}(i)$ represents maximum cooling release speed of ISTs during period i, kWh.

In addition, the remaining cooling capacity of ISTs at any given time can be expressed as follows:

$$Q_{\text{IST}}^{\text{re}}(i+1) = Q_{\text{IST}}^{\text{re}}(i) + Q_{\text{IST}}^{\text{ac}}(i) - Q_{\text{IST}}^{\text{cr}}(i), \tag{11}$$

where $Q_{\text{IST}}^{\text{re}}(i+1)$ represents the remaining cooling capacity of ISTs at the beginning of the period $(i+1)$, kWh.

Note that these calculations only present expressions for the various constraints to be called in the follow-up potential analysis step, while the specific data input and calculation will be done automatically by the computer. The call of the equations is further described in Section 2.8.2.

2.8. Energy-Saving Potential Analysis

2.8.1. Problem Definition and Principles of Potential Analysis

Generalized computational equations for energy and cost savings are presented in a companion paper [1]. For the present case with multiple different tariffs in a day (Table A1 in Appendix A), the energy cost savings for each day can be defined as follows:

$$\Delta J = J_{actual} - J_{benchmark} = \sum_{i=0}^{23} e_i (W_{actual}(i) - W_{benchmark}(i)), \tag{12}$$

where ΔJ represents the energy cost-saving of the air-conditioning system for one day, CNY; J_{actual} represents the actual energy costs of the air-conditioning system for one day, CNY; $J_{benchmark}$ is the benchmark energy cost of the air-conditioning system for one day, CNY; e_i is the electricity price for period i, CNY/kWh; $W_{actual}(i)$ is the actual energy consumption of the air-conditioning system for the period i, kWh; and $W_{benchmark}(i)$ is the benchmark energy consumption of the air-conditioning system for the period i, kWh.

As previously mentioned, only energy consumption and performance data of equipment in the chiller plant room are available concurrently in this case, lacking the necessary data related to cooling towers and AHUs. Therefore, this study focuses on the analysis and calculation of the saving potential of the chiller plant room, summarized as follows:

The benchmark energy cost is the optimization result of the system operation. Some principles need to be followed to calculate the benchmark value.

- For each cooling storage and release cycle, the operation of the air conditioning system shall be optimized and calculated according to the cooling load demand in each period and the constraints described in Section 2.7;
- For each period of a cooling storage and release cycle, the operation mode shall be determined according to the cooling load demand and electricity tariff;
- For each period determined as the operation mode of M4, it is necessary to further determine the respective cooling supply ratios of the ice and chillers;
- For each set of parameters resulting from the above principles, the total energy cost of the chiller plant room was calculated according to the results of the regression analysis in Section 2.6;
- Search for the minimum energy costs as the benchmark energy costs by continuously adjusting the operation mode and other parameters for each period during the cooling storage and release cycle.

The calculations were carried out under the following assumptions:

- The current remaining cooling capacity in the ISTs at the beginning of period 0 on the first day (i.e., 22 June 2013) is 0 kWh;
- The maximum cooling storage or release speed of ISTs is determined according to the performance curve provided by the manufacturer and the remaining cooling capacity in the ISTs at the beginning of the current period;
- The maximum accumulation of cooling or cooling release of ISTs during one cooling storage and release cycle is determined according to the performance information provided by the manufacturer.

2.8.2. Calculation of the Benchmark Energy Costs by Particle Swarm Optimization (PSO)

This study introduces the PSO algorithm to calculate the system's benchmark energy costs by considering the above principles and the global search and optimization capabilities of the algorithm.

As a branch of evolutionary computation, the PSO algorithm is a new swarm intelligence optimization algorithm proposed in 1995 by Kennedy and Eberhart [23]. The algorithm simulates the activity patterns of birds and fish flocks and achieves an optimal solution to the problem through inter-group individual collaboration. The particles search cooperatively in the region of feasible solutions. In addition to its own flight inertia,

each particle simultaneously draws on its own and the optimal global experience of the entire particle swarm to approach the optimal global solution. The PSO algorithm maintains a population of particles. The position of the particle denotes a feasible, if not the best, solution to the problem. The objective function value is improved by the optimum progress, which is required to change the particle position. The convergence condition always requires the input of the move iteration number of the particle. The moving rule for the particle's position can be depicted by the following equations [24]:

$$V_s(t+1) = wV_s(t) + C_1r_1(P_s - X_s(t)) + C_2r_2(G - X_s(t)), \tag{13}$$

$$X_s(t+1) = X_s(t) + V_s(t+1), \tag{14}$$

where Vs(*t*) represents the velocity vector of particle *s* in *t* time; $X_s(t)$ represents the position vector of particle *s* in *t* time; P_s is the personal best position of particle *s*; *G* is the best position of the particle found at the present time; *w* represents the inertia weight; C_1 and C_2 are two acceleration constants, called cognitive and social parameters, respectively; and r_1 and r_2 are two random functions in the range [0, 1].

Specifically, the solution that satisfies the actual case of this study can be obtained through the processes shown in Figure 2.
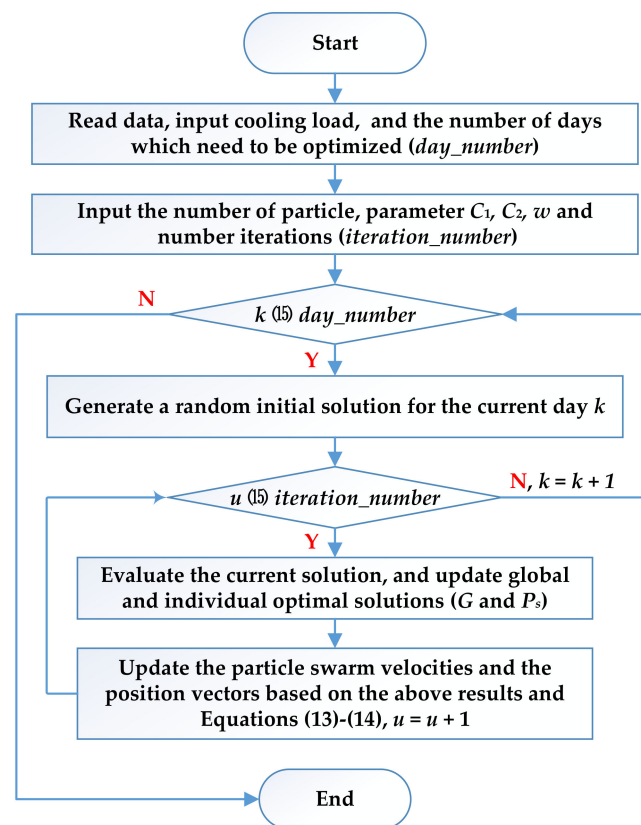
Figure 2. Flow chart of the algorithm for benchmark energy costs.

The algorithmic process in Figure 2 is illustrated in detail as follows:

- STEP 1: To start, read data from the dataset of the case system, and input the cooling load for each period *i* of each day and the number of days to be optimized (*day_number*) into the algorithm by computer:
- STEP 2: Manually input the number of particles, parameters C_1, C_2, *w*, and the number of iterations (*iteration_number*);
- STEP 3: Judge whether the current day *k* is smaller than *day_number*; If YES, enter the optimization process of day *k*, step forward; if NOT, the process ends;

- STEP 4: Generate a random initial solution for the current day k;
- STEP 5: If the iteration (u) for the current day k is smaller than *the iteration_number*, enter the particle swarm iteration cycle; otherwise, $k = k + 1$ and step forward to STEP 3;
- STEP 6: Evaluate the current solution and update the global and individual optimal solutions;
- STEP 7: Update the particle swarm velocities and the position vectors based on the results of the previous STEP 6 and in combination with Equations (13) and (14), $u = u + 1$, and step forward to STEP 5.

This study employs the above process for the day-to-day optimization of the calculations until the calculations for all the days are optimized and completed. Moreover, some key details are described as follows:

- Each solution for the current day k is represented by a position vector X_s;
- The evolution of the solution begins in the PSO with an initial solution consisting of initial particles;
- The initial solution is obtained by a random initial position of each particle; a matrix is employed for recording the operating modes and other status parameters of the case system;
- For the periods 0–8, one of the modes (M0 or M1) can be randomly selected. If M1 is selected, an accumulation of cooling capacity is generated. Call and make sure that Equations (6)–(8) and (11) are valid; otherwise, the mode should be reselected;
- For periods 9–23, one of the modes (M0, M2, M3, or M4) can be randomly selected. However, M0 should be selected as long as $Q_{CL}(i) = 0$. If M2 is selected, call and make sure that Equations (4) and (5) are both valid; otherwise, M4 is selected. If M3 is selected, call and make sure that Equations (4) and (9)–(11) are valid; otherwise, M4 is selected. When M4 is selected, the algorithm randomly generates the respective cooling supply ratio of the ice and chillers. Call and make sure that Equations (4), (5) and (9)–(11) are valid; otherwise, regenerate the respective cooling supply ratio.

3. Results and Discussion

3.1. Recognizing System Running Mode

This study employs the operation data of 1416 periods from 22 June to 19 August 2013 to model the decision tree using the CART algorithm to recognize the operating modes. The established decision tree is shown in Figure 3. In the modeling process, the $\Delta Gini$ values of the variables for each node containing multiple operating modes (also known as the internal node, i.e., nodes 0, 2, 3, and 5) are shown in Tables A3–A6 of Appendix A.
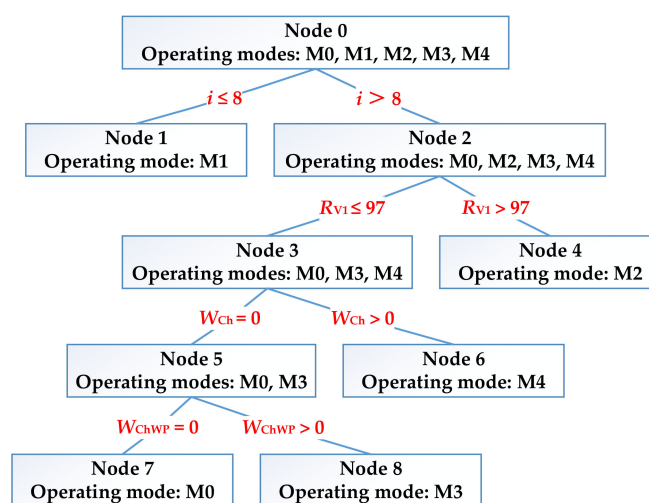

Figure 3. Decision-tree classification model for recognizing operation modes.

Among the four internal nodes of 0, 2, 3, and 5, in node 2, R_{V1} = 97 is determined as the splitting criterion based on the monitoring data for the opening of V1 (97.28 $\leq R_{V1} \leq$ 97.89 in the "on" state, while the maximum value of 61.20 in the "modulate" state). Meanwhile, the other three nodes use the theoretical values of the variable in the corresponding classifications to establish the criterion. This, in turn, enables the checking for anomalies in the actual operational data (e.g., the over-running in Section 3.4).

3.2. Regression Results of ME

Following the fitting method in Section 2.6, the fitting results and valid internals of energy consumption and flow rate for the nine pumps in the case system are presented in Table 4.

Table 4. Fitting results of energy-consumption and flow rate of the nine pumps in the case system.

| Items | Fitting Model | Estimates of Coefficient | | Range of Value | |
		β_0	β_1	W (kWh)	q (m^3/h)
ChWP-1	$W = \beta_1 (q + \beta_0)^3$	7.755	2.738×10^{-7}	[7.90, 13.50]	[298.94, 358.92]
ChWP-2	$W = \beta_1 (q + \beta_0)^3$	−18.501	3.475×10^{-7}	[7.90, 13.80]	[301.79, 359.68]
ChWP-3	$W = \beta_0 + \beta_1 q$	16.507	2.901×10^{-2}	[25.70, 27.00]	[316.94, 361.76]
GWP-1	$W = \beta_1 (q + \beta_0)^3$	136.140	7.399×10^{-8}	[2.00, 6.50]	[163.98, 308.41]
GWP-2	$W = \beta_1 (q + \beta_0)^3$	196.989	1.785×10^{-7}	[9.20, 9.80]	[175.14, 183.06]
GWP-3	$W = \beta_0 + \beta_1 q$	31.554	3.558×10^{-2}	[39.90, 41.30]	[234.58, 273.92]
CWP-1	$W = \beta_0 + \beta_1 q$	12.541	6.148×10^{-2}	[18.30, 26.60]	[93.67, 228.68]
CWP-2	$W = \beta_0 + \beta_1 q$	10.236	1.416×10^{-2}	[12.40, 12.70]	[152.76, 173.94]
CWP-3	$W = \beta_0 + \beta_1 q$	27.410	4.928×10^{-2}	[31.70, 38.40]	[87.04, 222.99]

Considering Ch-1 (Chiller #1) in the cooling mode (i.e., corresponding to operating modes M2 and M4) as an example, the fitting results of energy consumption based on Equation (3) in Section 2.6 are listed in Table 5. Similar results for the ice-building mode (i.e., corresponding to operating mode M1) of Ch-1 and two modes of Ch-2 (Chiller #2) are not listed because of space limitations.

Table 5. Fitting results of the energy-consumption for chiller 1(Ch-1) in cooling mode.

| Coefficient	Each Item of the Polynomial	Coefficient Estimate	Standard Error
β_0	1	−6630.420	1814.984
β_1	Q_{Ch}	−0.146	1.125
β_2	t_{cwChr}	181.475	50.063
β_3	Δt_{chwCh}	−84.588	303.087
β_4	Δt_{cwCh}	2,120.110	627.771
β_5	$(Q_{Ch})^2$	8.897×10^{-5}	0.001
β_6	$(t_{cwChr})^2$	−1.552	0.876
β_7	$(\Delta t_{chwCh})^2$	15.478	24.417
β_8	$(\Delta t_{cwCh})^2$	−145.624	84.335
β_9	$Q_{Ch} t_{cwChr}$	−0.020	0.030
β_{10}	$Q_{Ch} \Delta t_{chwCh}$	0.083	0.161
β_{11}	$Q_{Ch} \Delta t_{cwCh}$	−0.124	0.347
β_{12}	$t_{cwChr} \Delta t_{chwCh}$	1.965	7.730
β_{13}	$t_{cwChr} \Delta t_{cwCh}$	−14.364	10.743
β_{14}	$\Delta t_{chwCh} \Delta t_{cwCh}$	−46.076	60.213
β_{15}	t_{chwChs}	−6.715	54.304
β_{16}	$(t_{chwChs})^2$	0.129	0.827
β_{17}	$Q_{Ch} t_{chwChs}$	0.053	0.028
β_{18}	$t_{cwChr} t_{chwChs}$	−0.705	1.517
β_{19}	$\Delta t_{chwCh} t_{chwChs}$	1.184	6.318
β_{20}	$\Delta t_{cwCh} t_{chwChs}$	−10.991	14.230

3.3. Energy-Saving Potential Results

For the PSO algorithm, the values of C_1, C_2, and w may influence the computational results [24]. After loading the data for 59 d from the case system into the algorithm described in Section 2.8, better results can be obtained by multiple trial calculations to determine the parameters $C_1 = 2$, $C_2 = 2$, and $w = 0.6$. Table 6 shows the average results for 10 consecutive calculations at various iteration sizes. Moreover, the single minimum energy cost appears in the calculation process with an iteration scale of No. 6, 307,213.5 CNY. It saves 97,156.0 CNY or 24.03% compared to the actual energy cost of 404,369.5 CNY. The comparison between the two scenarios for each day is shown in Figure 4. The specific data in Figure 4 are listed in Table A7 in Appendix A, and the mean absolute percentage error (MAPE) of the optimal costs for each day are reported simultaneously in the table.

Table 6. Average results of various iteration sizes.

No.	Particle Quantity	Steps of Iterations	Time Cost (s)	Average Result (CNY)
1	50	50	273	326,987.4
2	50	100	541	326,311.8
3	50	200	1 080	324,513.9
4	100	50	544	313,138.6
5	100	80	866	309,388.9
6	100	100	1 083	308,598.6
7	100	150	1 610	311,192.5
8	100	200	2 151	310,063.8
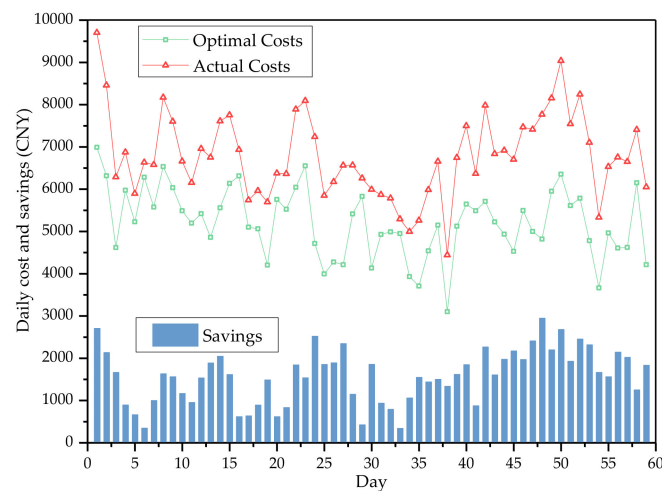
Figure 4. Comparison between the daily actual costs, optimal costs, and savings of the case system for 59 d from 22 June to 19 August 2013.

When combining these results, the following conclusions can be drawn about the systematic approach of this study:

First, in Figure 4, the case system has different levels of energy-saving potential for each day, indicating and validating the effectiveness of the optimization algorithm in this study. Second, it is evident from the results in Table 6 that the difference between the maximum and minimum values of the average results of eight types of iteration sizes is only 5.6%. This is in combination with the MAPE results of the optimization costs in Table A7 (ranging from 2.2% to 11.1%, with an average of 6.2%), demonstrating the high stability of the optimization algorithm. Third, owing to the complexity of the case system model and the numerous constraints, the process of the savings potential calculation is time-consuming. The algorithm needs to be further improved to address this shortcoming.

3.4. Discussion about Selection of Models

Considering that the main purpose of this paper was to implement and validate the systematic approach by applying it to a real-life system, we did not focus on the specific models and algorithms used for the case study. It is noteworthy that the specific models applied in this study exhibit considerable necessity and superiority.

In terms of the model used for recognition of system-operation mode, the CART algorithm was chosen to construct the decision tree model in this paper. Relative to other clustering models, the CART has better interactivity so that the background knowledge for the decision tree can be incorporated. [20] Furthermore, the actual and theoretical values of the relevant variables during system operation were both considered in the splitting criterion. For example, when V1 is in the "on" state, the actual value of R_{V1} was between 97.28 and 97.89, instead of the theoretical value of 100. Considering the real-life operation, this study determined $R_{V1} = 97$ as the splitting criterion of node 2. As for the splitting criterion of equipment energy consumption, this paper took the theoretical values in Table 3. The model was thus built not only to identify the system-operation mode but also eligible to check for anomalies in the actual operating data. For instance, according to the model (Figure 3), the moment in Table A2 in Appendix A at 09:00 on 18 August 2013 was identified as M3 (the relevant parameters for this moment were: $i = 9$, $R_{V1} = 1.44$, $W_{Ch} = 0$, $W_{ChWP} = 7.9$). Therefore, based on the rules shown in Table 3, the condensing water pumps should be in a shutdown state, and W_{CWP} should be 0, while the actual value of W_{CWP} was 1.1. It indicates an abnormal running of the pumps, where the energy-saving potential exists. After statistics, the energy costs caused by over-running account for 3.8% of the total costs. The detection of such anomalies would further demonstrate the superiority of the model used in this study.

With respect to the optimization model, indeed, many optimization algorithms can be applied to the potential analysis session. In addition to the PSO method, we also examined the performance of other algorithms, including genetic algorithm (GA) and ant colony optimization (ACO). As seen in Figure A1, PSO had better optimization results at the same time-consuming level (~1000 s) relative to GA and ACO—15% and 8% more energy-saving potential were obtained by PSO than GA and ACO, respectively. We also acknowledge that comparative studies of the different optimization algorithms in terms of feasibility and consumption need to be undertaken in the future to obtain the greatest possible energy–cost-saving potential.

3.5. Advantages of the Systematic Method

The detailed case study verifies and demonstrates the feasibility, effectiveness, and stability of the systematic approach. In addition, the proposed method exhibits obvious advantages over the conventional OTI method. Owing to the elimination of a significant amount of on-site work, including communication with users and system testing in the OTI method, the proposed method can reduce the time consumed for energy-saving potential analysis of complex systems from weeks or months to days, or even shorter. Likewise, the proposed method consumes much fewer human resources than the conventional method. In addition, the conventional OTI method focuses on more specific problems. Hence, the follow-up measurements and analyses are mainly focused on achieving the local optimization of specific equipment in the system. The results may work well for the specific equipment, but not necessarily for the system. On the contrary, the proposed method is more comprehensive, considers the extensive constraints in the actual operation of the system, and utilizes artificial intelligence algorithms to achieve the global optimization of the entire system. Furthermore, the systematic method allows for the modular and batch processing of data, as well as remote analysis and online diagnosis, thus providing better versatility and scalability.

3.6. Limitations and Future Outlook

The limitations of the proposed systematic method and the case study need to be alleviated in future work and are listed as follows:

- As previously stated, owing to the lack of sufficient data on cooling towers and terminal AHUs for analysis, the case study could not optimize the energy costs simultaneously in the final optimization. However, this issue does not affect the scientific and systematic nature of the study;
- The final optimization of the method pertained to the whole system rather than the individual devices. This limits the optimization solutions based on the performance of individual devices.
- It is crucial to achieve the energy-saving operation of the air-conditioning system through the improvement of the control strategy, to attain or approach the result of energy-saving potential calculation. Obtaining a set of analysis and operation methods combining potential calculation and optimization control is a practical problem that needs further attention;
- In the research process of diagnostic methods and the employment of these methods to diagnose real air-conditioning systems, the visualization and professional interaction of information and data cannot only improve work efficiency but also have a significant impact on the understanding and application of the relevant results. This is a potential direction for future research.

4. Conclusions

The paper presents a detailed case study of an ice-storage air-conditioning system using the method based on data-mining proposed in the companion paper. The raw data were preprocessed prior to recognizing the constant- and variable-speed devices in the system. The classification- and regression-tree algorithms were used to identify the operating modes of the system. The regression models between the energy-consumption and operating-state parameters of the nine pumps and two chillers were fitted. Subsequently, the constraints related to the system operation were summarized. From the results, the particle swarm optimization method was applied to obtain the benchmark energy cost and the consequent cost-saving potential. The cost-saving potential for the chiller plant room during the 59 d of investigation reached as high as 24.03%. The case study validates and demonstrates the feasibility, effectiveness, and stability of the systematic approach.

Compared with conventional methods, which take weeks or even longer for energy-saving potential analyses, the proposed method takes only a few days or less. This enhanced speed indicates that the new systematic approach effectively identifies system defaults and reduces the time spent on troubleshooting. The proposed method provides a new approach for studying actual operation data, which is significant in enhancing the energy efficiency of air-conditioning systems. Future research is warranted to develop corresponding control strategies based on the potential analysis results, investigate better optimization algorithms, and visualize the analysis process.

Author Contributions: Conceptualization and project administration, R.M.; methodology and data analysis, R.M. and S.Y.; academic guidance, resources, and supervision, X.W., X.-C.W., N.Y., X.Y., and M.S.; writing—original draft, R.M., and S.Y.; writing—review and editing, X.W., X.-C.W., M.S., X.Y., N.Y., and S.Y.; funding acquisition, X.Y., M.S., and R.M. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Abbreviations

ACO	Ant colony optimization
AHU	Air handling unit
ANN	Artificial neural network
CART	Classification and regression trees
Ch	Chiller
ChWP	Chilled-water pump
ChPR	Chiller plant room
CWP	Condensing-water pump
GA	Genetic algorithm
GWP	Glycol water pump
HVAC	Heating, ventilating and air conditioning
IST	Ice storage tank
MAPE	Mean absolute percentage error
NN	Neural network
OTI	Observation/question, test/calculation, and identification/resolution
PHEX	Plate heat exchanger
PSO	Particle swarm optimization
RSM	Response surface methodology
SAX	Symbolic aggregate approximation
SI	Système international
SVM	Support vector machine
USRT	US refrigeration ton

Appendix A

Table A1. Electricity tariffs (e_i) at different time intervals of the day.

i	Time Interval	e_i (CNY/kWh)
0	23:00–00:00	0.2884
1	00:00–01:00	0.2884
2	01:00–02:00	0.2884
3	02:00–03:00	0.2884
4	03:00–04:00	0.2884
5	04:00–05:00	0.2884
6	05:00–06:00	0.2884
7	06:00–07:00	0.2884
8	07:00–08:00	0.8844
9	08:00–09:00	0.8844
10	09:00–10:00	1.1644
11	10:00–11:00	1.1644
12	11:00–12:00	1.0244
13	12:00–13:00	0.8844
14	13:00–14:00	0.8844
15	14:00–15:00	1.1644
16	15:00–16:00	1.1644
17	16:00–17:00	1.0244
18	17:00–18:00	0.8844
19	18:00–19:00	0.8844
20	19:00–20:00	1.1644
21	20:00–21:00	1.1644
22	21:00–22:00	0.8844
23	22:00–23:00	0.8844

Table A2. Selected raw data collected on 18 August 2013.

Date Time	t_{OA}	t_{cws}	t_{cwr}	t_{chws}	t_{chwr}	$t_{chwCh-1s}$	$t_{chwCh-1r}$	$t_{cwCh-1s}$	$t_{cwCh-1r}$	t_{gwISTr}	t_{gwISTs}	R_{V1}	q_{chw}	Q_{CL}	m_{isIST}	W_{CWP-1}	W_{CWP-2}	W_{CWP-3}	W_{GWP-1}	W_{GWP-2}	W_{GWP-3}	W_{ChWP-1}	W_{ChWP-2}	W_{ChWP-3}	W_{Ch-1}	W_{Ch-2}	W_{ChPR}
18 August 2013 00:00:00	28.35	35.37	30.60	19.36	19.22	5.27	5.34	35.18	30.52	2.40	6.21	1.44	0.17	−0.03	635.42	24.8	6.6	6.3	0	26.4	40.8	0	0	0	261.8	257	623.7
18 August 2013 01:00:00	28.35	33.32	30.60	19.36	19.42	−0.16	−0.16	33.25	30.54	−2.74	0.81	1.44	0.16	0.01	623.26	24	12	0	0	26.4	40.8	0	0	0	265.6	260	628.8
18 August 2013 02:00:00	28.25	32.82	30.30	19.46	19.52	−3.33	−3.35	32.79	30.18	−5.74	−2.36	1.44	0.16	0.01	689.52	23.9	11.9	0	0	26.4	40.9	0	0	0	266.3	252.9	622.3
18 August 2013 03:00:00	27.75	32.91	30.20	19.56	19.72	−2.25	−1.91	32.91	30.24	−4.57	−1.02	1.44	0.16	0.03	1626.46	23.9	11.8	0	0	26.5	40.9	0	0	0	266.8	259.8	629.7
18 August 2013 04:00:00	27.95	32.91	30.30	19.56	19.82	−2.54	−2.27	32.96	30.30	−4.77	−1.36	1.44	0.16	0.05	2373.00	23.9	11.9	0	0	26.4	41	0	0	0	266.8	254.4	624.4
18 August 2013 05:00:00	28.05	32.91	30.30	19.66	20.02	−2.79	−2.51	32.96	30.30	−4.98	−1.63	1.44	0.16	0.07	3101.35	24	11.9	0	0	26.4	40.9	0	0	0	267.3	251.3	621.8
18 August 2013 06:00:00	28.05	33.01	30.41	19.66	20.12	−3.03	−2.81	33.02	30.42	−5.27	−1.88	1.44	0.16	0.09	3831.61	23.9	11.9	0	0	26.5	40.9	0	0	0	267.9	253.8	624.9
18 August 2013 07:00:00	28.15	32.71	30.20	19.76	20.32	−3.21	−2.99	32.72	30.12	−5.47	−2.10	1.44	0.16	0.11	4558.58	9.9	4.9	0	0	10.4	16.2	0	0	0	101.4	93.5	236.3
18 August 2013 08:00:00	28.25	30.29	30.30	19.76	20.42	1.07	−2.20	30.36	29.70	−2.07	−1.04	1.44	0.16	0.13	4574.85	0	0	0	0	0	0	0	0	0	0	0	0
18 August 2013 09:00:00	28.75	30.29	30.30	19.86	20.51	2.32	−1.60	30.30	29.16	−1.27	−0.39	1.44	0.16	0.13	4542.40	0	0	1.1	2.4	0	0	0	7.9	0	0	0	11.4
18 August 2013 10:00:00	29.25	30.09	30.30	12.73	16.78	2.86	−0.87	30.06	28.86	15.73	4.46	23.10	86.37	409.84	3457.71	0	0	0	3.2	0	0	0	12.1	0	0	0	15.3
18 August 2013 11:00:00	27.55	30.09	30.30	12.04	16.45	3.82	−0.21	29.82	28.62	15.37	5.38	16.48	92.78	488.93	2515.13	0.7	0	0	3.3	0	1.1	0	12.2	1	1.2	0	19.5
18 August 2013 12:00:00	27.55	32.84	29.24	14.84	17.75	13.73	16.88	33.38	29.21	14.43	6.22	97.57	94.33	329.83	1807.82	25.9	0	0	1.5	0	40.6	9.7	0	6.8	262.2	17.2	363.9
18 August 2013 13:00:00	27.54	33.34	29.25	15.04	18.06	14.27	17.18	33.42	29.21	14.62	6.97	97.57	96.38	347.88	1821.17	25.4	0	0	0	0	40.8	13	0	0	272	0	351.2
18 August 2013 14:00:00	27.85	33.53	29.54	15.24	18.25	14.57	17.42	33.66	29.57	14.72	7.69	97.57	96.49	346.75	1843.90	10.6	0	0	2.7	0	16.2	6.2	7.1	0	103.1	0	145.9
18 August 2013 15:00:00	28.25	30.76	29.64	10.23	15.16	18.99	17.51	30.00	29.39	13.59	7.01	1.52	92.02	546.69	1021.52	0	0	0	9.4	0	0	0	12.5	0	0	0	21.9
18 August 2013 16:00:00	28.55	30.66	29.64	12.15	15.98	19.65	17.15	30.00	28.79	14.61	10.22	1.52	88.76	405.65	702.72	12.1	0	0	5.1	0	19	0	5.7	14.9	96.2	0	153
18 August 2013 17:00:00	28.95	33.83	30.02	15.63	18.69	14.90	17.85	34.14	30.03	15.40	10.33	97.44	94.69	348.18	690.35	25.7	0	0	1.3	0	40.5	0.3	0	26.8	262.5	14.2	371.3
18 August 2013 17:59:00	28.95	34.23	30.33	16.35	19.29	15.57	18.51	34.46	30.33	15.50	10.74	97.44	95.95	338.33	700.05	25.1	0	0	0	0	40	0	0	26	263.3	0	354.4
18 August 2013 18:00:00	28.95	34.23	30.33	16.35	19.29	15.57	18.51	34.46	30.27	15.50	10.74	97.44	97.50	343.80	700.05	25.5	0	0	0	0	40.7	0	0	26.4	267.8	0	360.4
18 August 2013 19:00:00	28.55	33.47	30.13	16.89	19.39	18.04	19.29	30.69	30.09	15.50	11.31	97.44	103.95	311.54	709.61	26.1	0	0	5	0	39.3	1.6	0	25.7	254.2	122.6	474.5
18 August 2013 20:00:00	28.24	37.26	31.31	12.67	17.25	12.62	15.07	37.05	31.39	15.50	12.11	97.44	100.74	552.70	725.60	26.6	0	0	8.5	0	38.4	1.8	0	26.2	264.2	266.4	632.1
18 August 2013 21:00:00	28.14	36.96	31.21	11.67	16.24	11.74	14.05	36.70	31.21	15.59	12.33	97.44	97.30	533.77	736.38	26.5	0	0	8.6	0	38.3	1.8	0	26.3	264.1	266.9	632.5
18 August 2013 22:00:00	28.14	36.76	31.21	10.87	15.34	10.97	13.20	36.64	31.15	15.79	11.83	97.44	96.02	514.53	737.41	19.1	0	0	6	0	27	1.1	0	18.9	182.1	188.4	442.6
18 August 2013 23:00:00	28.04	31.65	31.21	13.63	13.58	14.56	13.16	31.48	30.85	16.09	11.05	1.55	0.16	−0.01	713.68	14	0	0	0	9	22.4	0	0	0	133.2	52.3	230.9

Table A3. $\Delta Gini$ values for variables at node 0.

Variables	$\Delta Gini$
i	0.299
W_{AHU}	0.276
W_{ChWP}	0.270
W_{GWP}	0.190
R_{V1}	0.188
W_{Ch}	0.172
W_{CWP}	0.128

Table A4. $\Delta Gini$ values for variables at node 2.

Variables	$\Delta Gini$
R_{V1}	0.172
W_{Ch}	0.146
W_{GWP}	0.137
W_{ChWP}	0.127
W_{CWP}	0.076
W_{AHU}	0.056
i	0.054

Table A5. $\Delta Gini$ values for variables at node 3.

Variables	$\Delta Gini$
W_{Ch}	0.155
W_{GWP}	0.146
W_{ChWP}	0.140
W_{CWP}	0.074
W_{AHU}	0.058
R_{V1}	0.054
i	0.046

Table A6. $\Delta Gini$ values for variables at node 5.

| Variables	$\Delta Gini$
W_{ChWP}	0.066
W_{GWP}	0.063
i	0.063
W_{AHU}	0.060
R_{V1}	0.029
W_{CWP}	0.009
W_{Ch}	0.000

Table A7. Daily actual costs, optimal costs and their mean absolute percentage error, and savings of the case system for 59 d from 22 June to 19 August 2013.

Day	Actual Costs (CNY)	Optimal Costs (CNY)	Savings (CNY)	MAPE [1] of Optimal Costs (%)
1	9700.0	6989.1	2710.9	2.2%
2	8454.8	6314.7	2140.0	3.2%
3	6285.1	4617.6	1667.5	7.9%
4	6872.3	5972.0	900.3	5.5%
5	5893.8	5225.0	668.8	6.6%

Table A7. *Cont.*

Day	Actual Costs (CNY)	Optimal Costs (CNY)	Savings (CNY)	MAPE [1] of Optimal Costs (%)
6	6630.1	6278.8	351.3	3.4%
7	6576.5	5573.7	1002.9	2.6%
8	8170.0	6532.8	1637.1	5.3%
9	7601.2	6034.0	1567.3	5.0%
10	6655.7	5487.0	1168.7	4.7%
11	6151.9	5194.5	957.4	5.8%
12	6955.4	5418.4	1537.0	8.3%
13	6749.7	4858.5	1891.2	9.7%
14	7610.0	5557.6	2052.4	5.1%
15	7752.2	6134.0	1618.2	5.1%
16	6933.4	6312.6	620.8	5.5%
17	5738.6	5097.4	641.1	9.2%
18	5959.4	5061.3	898.1	6.9%
19	5695.0	4202.9	1492.1	11.1%
20	6376.4	5754.5	621.9	6.0%
21	6360.2	5521.1	839.1	6.2%
22	7886.9	6041.3	1845.6	4.7%
23	8089.9	6548.6	1541.3	3.6%
24	7236.2	4712.6	2523.6	4.0%
25	5848.2	3991.8	1856.4	8.7%
26	6171.7	4275.7	1896.0	8.4%
27	6561.4	4211.0	2350.4	7.8%
28	6565.0	5413.0	1152.0	3.0%
29	6257.1	5827.2	429.9	5.0%
30	5990.5	4131.1	1859.4	7.7%
31	5866.5	4926.2	940.3	7.1%
32	5784.7	4989.7	795.0	6.2%
33	5290.7	4946.4	344.3	7.8%
34	4995.0	3929.5	1065.5	6.5%
35	5258.5	3706.8	1551.8	8.8%
36	5983.7	4539.1	1444.6	5.4%
37	6654.7	5147.4	1507.4	5.5%
38	4439.0	3099.5	1339.4	8.8%
39	6745.4	5122.6	1622.8	4.9%
40	7495.8	5644.6	1851.2	6.8%
41	6366.0	5487.9	878.1	5.1%
42	7981.3	5707.8	2273.5	3.9%
43	6833.6	5224.4	1609.3	5.9%
44	6915.0	4934.6	1980.4	5.9%
45	6701.3	4525.8	2175.5	4.6%
46	7467.3	5494.3	1973.0	6.7%
47	7411.0	4997.0	2413.9	7.0%
48	7766.7	4815.9	2950.9	7.9%
49	8151.1	5949.6	2201.4	2.6%
50	9035.6	6352.2	2683.4	3.7%
51	7541.6	5608.4	1933.2	7.1%
52	8242.6	5783.4	2459.2	3.4%
53	7102.0	4780.8	2321.2	7.9%
54	5329.8	3661.2	1668.5	10.1%
55	6528.5	4962.1	1566.4	8.7%
56	6751.7	4605.2	2146.5	4.8%
57	6647.2	4621.1	2026.1	7.4%
58	7406.0	6149.2	1256.8	4.9%
59	6046.7	4211.0	1835.7	10.2%

[1] MAPE: mean absolute percentage error.
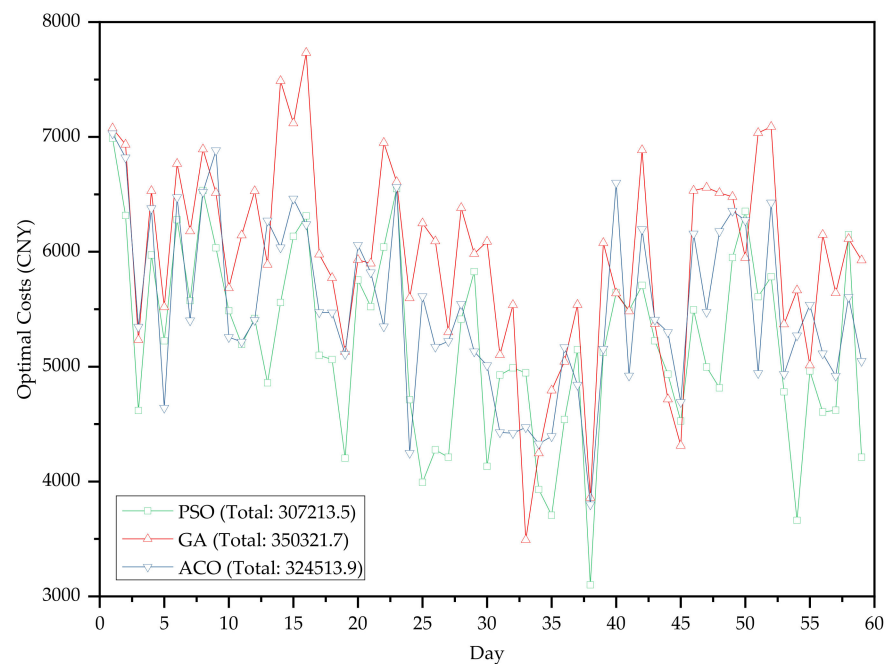
Figure A1. Comparison of optimal costs of the case system calculated by three algorithms (particle swarm optimization (PSO), genetic algorithm (GA) and ant colony optimization (ACO)) for 59 d from 22 June to 19 August 2013.

References

1. Ma, R.; Wang, X.; Wang, X.-C.; Shan, M.; Yang, X.; Yu, N.; Yang, S. Systematic method for energy saving potential calculation of air conditioning systems via data mining. Part I: Methodology. *Energies* **2020**, *14*, 81. [CrossRef]
2. Zhao, Y.; Zhang, C.; Zhang, Y.; Wang, Z.; Li, J. A review of data mining technologies in building energy systems: Load prediction, pattern identification, fault detection and diagnosis. *Energy Built Environ.* **2020**, *1*, 149–164. [CrossRef]
3. Zhou, Q.; Wang, S.; Xiao, F. A novel strategy for the fault detection and diagnosis of centrifugal chiller systems. *HVAC R Res.* **2009**, *15*, 57–75. [CrossRef]
4. Han, H.; Gu, B.; Kang, J.; Li, Z.R. Study on a hybrid SVM model for chiller FDD applications. *Appl. Therm. Eng.* **2011**, *31*, 582–592. [CrossRef]
5. Van Every, P.M.; Rodriguez, M.; Jones, C.B.; Mammoli, A.A.; Martínez-Ramón, M. Advanced detection of HVAC faults using unsupervised SVM novelty detection and Gaussian process models. *Energy Build.* **2017**, *149*, 216–224. [CrossRef]
6. Yan, R.; Ma, Z.; Zhao, Y.; Kokogiannakis, G. A decision tree based data-driven diagnostic strategy for air handling units. *Energy Build.* **2016**, *133*, 37–45. [CrossRef]
7. Lee, W.Y.; Park, C.; House, J.M.; Kelly, G.E. Fault diagnosis of an air-handling unit using artificial neural networks. *ASHRAE Trans.* **1996**, *102*, 540–549.
8. Magoulès, F.; Zhao, H.; Elizondo, D. Development of an RDP neural network for building energy consumption fault detection and diagnosis. *Energy Build.* **2013**, *62*, 133–138. [CrossRef]
9. Patnaik, D.; Marwah, M.; Sharma, R.; Ramakrishnan, N. Sustainable operation and management of data center chillers using temporal data mining. In Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Paris, France, 28 June–1 July 2009; pp. 1305–1314.
10. Chen, Y.; Wen, J. Whole building system fault detection based on weather pattern matching and PCA method. In Proceedings of the 2017 3rd IEEE International Conference on Control Science and Systems Engineering (ICCSSE), Beijing, China, 17–19 August 2017; pp. 728–732.
11. Li, G.; Hu, Y.; Chen, H.; Li, H.; Hu, M.; Guo, Y.; Liu, J.; Sun, S.; Sun, M. Data partitioning and association mining for identifying VRF energy consumption patterns under various part loads and refrigerant charge conditions. *Appl. Energy* **2017**, *185*, 846–861. [CrossRef]
12. Li, Z.X.; Renault, F.L.; Gómez, A.O.C.; Sarafraz, M.M.; Khan, H.; Safaei, M.R.; Filho, E.P.B. Nanofluids as secondary fluid in the refrigeration system: Experimental data, regression, ANFIS, and NN modeling. *Int. J. Heat Mass Transf.* **2019**, *144*, 118635. [CrossRef]
13. Sarafraz, M.M.; Tlili, I.; Tian, Z.; Bakouri, M.; Safaei, M.R. Smart optimization of a thermosyphon heat pipe for an evacuated tube solar collector using response surface methodology (RSM). *Phys. A Stat. Mech. its Appl.* **2019**, *534*, 122146. [CrossRef]

14. Sarafraz, M.M.; Safaei, M.R.; Goodarzi, M.; Arjomandi, M. Experimental investigation and performance optimisation of a catalytic reforming micro-reactor using response surface methodology. *Energy Convers. Manag.* **2019**, *199*, 111983. [CrossRef]

15. Rosas-Flores, J.A.; Rosas-Flores, D. Potential energy savings and mitigation of emissions by insulation for residential buildings in Mexico. *Energy Build.* **2020**, *209*, 109698. [CrossRef]

16. Zeng, Y.; Zhang, Z.; Kusiak, A. Predictive modeling and optimization of a multi-zone HVAC system with data mining and firefly algorithms. *Energy* **2015**, *86*, 393–402. [CrossRef]

17. Kwame, A.B.O.; Troy, N.V.; Hamidreza, N. A Multi-Facet Retrofit Approach to Improve Energy Efficiency of Existing Class of Single-Family Residential Buildings in Hot-Humid Climate Zones. *Energies* **2020**, *13*, 1178. [CrossRef]

18. Tian, Z.; Si, B.; Wu, Y.; Zhou, X.; Shi, X. Multi-objective optimization model predictive dispatch precooling and ceiling fans in office buildings under different summer weather conditions. *Build. Simul.* **2019**, *12*, 999–1012. [CrossRef]

19. Ma, R.; Wang, X.; Shan, M.; Yu, N.; Yang, S. Recognition of variable-speed equipment in an air-conditioning system using numerical analysis of energy-consumption data. *Energies* **2020**, *13*, 4975. [CrossRef]

20. Han, J.; Kamber, M.; Pei, J. *Data Mining: Concepts and Techniques*, 3rd ed.; Elsevier Inc.: Waltham, MA, USA, 2012; ISBN 9780123814791.

21. Zhang, G.; Chen, L. Opmization of energy consumption in chilled plant. *Fluid Mach.* **2012**, *40*, 75–80.

22. Ma, R. The Study of Energy Efficiency Diagnosis Methodology for Air-Conditioning Systems with the Actual Operating Data. Ph.D. Thesis, Southwest Jiaotong University, Chengdu, China, 2014.

23. Kennedy, J.; Eberhart, R. Particle Swarm Optimization. In Proceedings of the IEEE International Conference on Neural Networks, Perth, Australia, 27 November–1 December 1995; Volume 4, pp. 1942–1948.

24. Ma, R.-J.; Yu, N.-Y.; Hu, J.-Y. Application of particle swarm optimization algorithm in the heating system planning problem. *Sci. World J.* **2013**, *2013*, 718345. [CrossRef] [PubMed]