Recognition of Variable-Speed Equipment in an Air-Conditioning System Using Numerical Analysis of Energy-Consumption Data

Rongjiang Ma [1,2], Xianlin Wang [1,3,4], Ming Shan [1], Nanyang Yu [2] and Shen 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.)
[2]   School of Mechanical Engineering, Southwest Jiaotong University, Chengdu 610031, China; rhinos@126.com
[3]   State Key Laboratory of Air-conditioning Equipment and System Energy Conservation, Zhuhai 519070, China
[4]   Gree Electric Appliances, Inc. of Zhuhai, Zhuhai 519070, China
[*]   Correspondence: yangshen13@tsinghua.org.cn

Abstract: Motor-driven equipment (ME) is one of the key components in an air-conditioning system, which contributes to the vast majority of the total energy consumption by air-conditioning systems. Distinguishing variable- and constant-speed equipment is important since the energy simulation models of the two types differ. Traditionally, types of ME are known in advance, and energy consumption data are consequently analyzed. However, in the application scenarios of energy consumption data mining, precedent information on the ME type could be missing. Thus, this study applies this process in reverse, providing new insight into energy consumption data of ME to recognize variable-speed ME in an air-conditioning system. The energy consumption data of ME in an air-conditioning system implemented in a commercial building were collected and numerically analyzed. A proposed simple parameter, coefficient of the median, and several numerical parameters were calculated and used to distinguish variable- from constant-speed ME. Results showed that the energy consumption data distributions of the two types of ME differed. The proposed coefficient of the median could successfully distinguish variable- from constant-speed ME, and it could be applied as an important step in energy consumption data mining of air-conditioning systems.

Keywords: motor-driven equipment (ME); air conditioning; numerical analysis; variable speed; data mining

1. Introduction

Air-conditioning systems are major sources of energy consumption in buildings [1,2]. Therefore, energy conservation for air-conditioning systems is of primary importance in building energy saving. In an air-conditioning system, the operation of various types of equipment is driven by motors, such as electric chillers, fans, and pumps. Therefore, motor-driven equipment (ME) is a key component in air-conditioning systems, especially in fluid transport subsystems. These subsystems transport refrigerant, cooled/heated air, chilled water, condensing water, or even a glycol solution to ensure the working cycle of the system. The energy consumed by ME makes up the vast majority of the total energy consumption of air-conditioning systems [3]. Thus, reducing the energy costs of ME may be an important approach to energy conservation for air-conditioning systems. Variable-speed ME has been used in air-conditioning systems for more than 40 years [4]. By changing the working frequency corresponding to flow-rate variations, variable-speed ME can generally cost less energy than constant-speed ME under the same conditions [5–7].

For air-conditioning systems equipped with both variable- and constant-speed ME, recognition of each equipment type is vital for energy-consumption analysis, particularly for energy simulations of ME, since the models of variable- and constant-speed ME are different [8–10]. In previous studies, the type of ME was known in advance, and the energy consumption data of the ME were then analyzed on the basis of known conditions [11–13]. For instance, several studies focused on the significant benefits brought about by the application of variable-speed drive technology to various types of ME of air-conditioning systems. For example, studies achieved comparative analyses of the energy performance of variable-speed compressors [4,14], pumps [15–17], fans [18,19], chiller plants [20–22], and even entire systems [23,24] against the performance of on/off-controlled constant-speed ME. However, what about going in reverse?

Normally, variable/constant-speed ME of an air-conditioning system is clarified by the designer during its design process. However, with the development of information technology and easy access to operating data, data-mining research through the use of actual operating data is now an important and effective tool for energy-saving diagnosis and operational optimization for air-conditioning systems [25]. In the context of automatic data mining of massive energy consumption data for energy-saving diagnosis [26–28], the codes normally have no such precedent information [29]. Instead, all the information is expected to come from the data-mining analysis of actual energy consumption data. It is, therefore, of great interest to investigate whether the energy consumption data of ME can reveal the characteristics and thus determine the type of ME. In addition, if the type of ME is already known, such an investigation could also check whether the variable-speed ME is working properly. Studies previously demonstrated the feasibility of using the intrinsic correlations of operational data to infer the operating state of air-conditioning systems [30]. However, more studies focused on the diagnosis of system faults using different methods, such as neural networks [31,32], wavelet analysis [33,34], principal component analysis [35,36], and combinatorial methods [37,38], or on making predictions related to future data by using existing data [39,40]. The few analyses of variable-speed systems were also based on the premise that it is known in advance whether it is a variable-speed device or not [41,42]. The paucity of literature examining the above issue prompted us to carry out this study.

This study aims to provide new insights into energy consumption data of ME, applied to the recognition of variable-speed ME in an air-conditioning system. The energy-consumption data of ME in an air-conditioning system operating in a commercial building were collected and numerically analyzed. A simple indicator parameter proposed by the study, coefficient of the median, and several numerical parameters were calculated and used to distinguish variable- from constant-speed ME. The results deepen our understanding of energy consumption characteristics related to the ME, and they provide insight into an important aspect of energy consumption-data mining in air-conditioning systems.

2. Methodology

2.1. ME in the Studied Air-Conditioning System

The studied air-conditioning system was installed in a five-story commercial building with an area of 30,000 m^2 located in Shenzhen, China. The system was coupled with an ice-storage system in order to take advantage of the lower electricity price at night. As shown in Figure 1, the system included two constant-speed chillers (Ch-1 and Ch-2) of the same brand and model, five sets of variable-speed air-handling units (AHUs; AHU-1, AHU-2, AHU-3, AHU-4, and AHU-5) of the same brand but different models, and nine pumps of the same brand but different models. The chillers, AHUs, and pumps were each of different brands, all of which were newly set up in the same year and operated properly for three years up to the time of data collection for this study. Furthermore, the nine pumps included three condensing-water pumps (CWP-1, CWP-2, and CWP-3) to achieve a condensing-water circle between chiller and cooling tower, three chilled-water pumps (ChWP-1, ChWP-2, and ChWP-3) to guarantee a chilled-water loop between heat exchanger and air-handling units at the terminal, and three glycol water pumps (GWP-1, GWP-2, and GWP-3) to transport the

glycol solution between any two or all three of the chiller, the ice-storage tank, and the heat exchanger. Four of the nine pumps could operate with variable speed, while the other five worked with constant speed. However, before this study, the authors were not aware of which pumps were variable-speed ME. This was later established on the basis of energy consumption data mining.
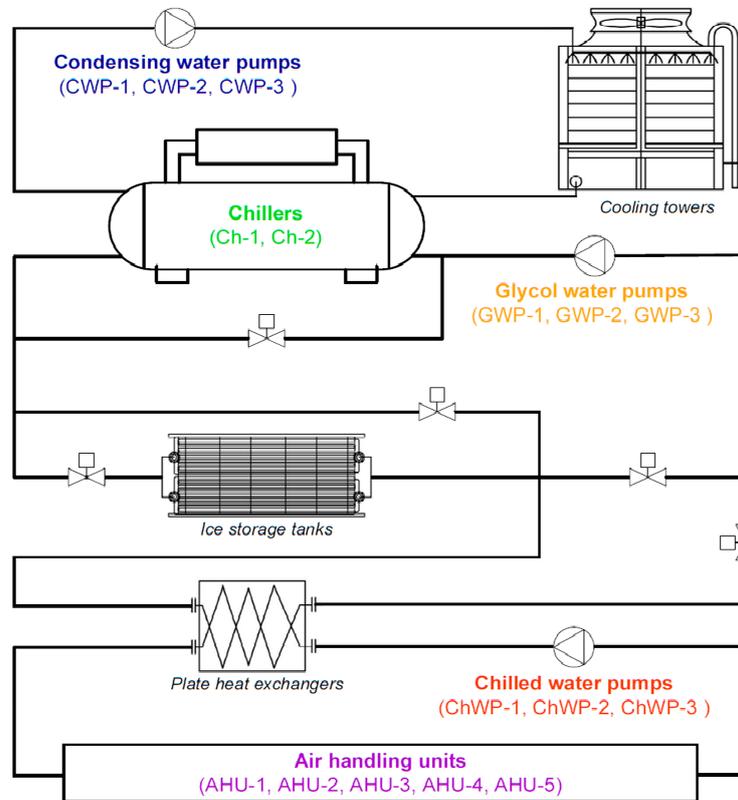
Figure 1. Schematic of studied air-conditioning system.

2.2. Collection and Processing of ME Energy Consumption Data

A comprehensive energy consumption monitoring platform was established for the air-conditioning system to monitor the electricity consumption of each component of the system. However, due to logistics and other constraints, the energy consumption of cooling towers was not available in the present case. Therefore, a follow-up study was employed to determine the energy consumption data of the remaining ME, including the aforementioned nine pumps, five AHUs, and two chillers. As it was expected that this study would be a true reflection of the use of the ME, data collection was conducted without any planned interventions or controls. For the involved ME, the electricity consumption of each device was recorded every hour from 19 May to 20 August 2013. Lastly, 2262 electricity consumption items were obtained for each device.

The raw data of electricity consumption of the ME and the corresponding time downloaded from the platform were then processed before numerical analysis. First, data with zero electricity consumption were eliminated since they demonstrated that the ME was turned off at the time. Second, data with duplicated time records were deleted to avoid repeated counts. Lastly, a processed electricity consumption data list with a time record was obtained for each device.

2.3. Numerical Analysis of Energy Consumption Data

Data distribution analysis can reveal the characteristics of energy consumption data. Nevertheless, using data distribution fitting can lead to extremely heavy computational loads, particularly with massive amounts of data. Instead, the numerical characteristics of data are important for data

statistics, as they can intuitively reveal the data distribution and can easily be obtained. In this study, some commonly used parameters were introduced, including dimensional parameters such as maximal value E_{max}, minimal value E_{min}, range E_{range}, mean value E_{mean}, median value E_{median}, mode E_{mode}, variance E_{var}, and standard deviation E_{sd}, and dimensionless parameters, including coefficient of variation C_v, coefficient of skewness C_s, and coefficient of kurtosis C_k.

Specifically, coefficient of variation C_v is defined as the ratio of standard deviation to the mean value of data [43], calculated using Equation (1).

$$C_v = \frac{E_{sd}}{E_{mean}}. \tag{1}$$

Coefficient of skewness C_s indicates the asymmetry of data distribution [44], calculated using Equation (2).

$$C_s = \frac{\frac{1}{n}\sum\limits_{i=1}^{n}(E_i - E_{mean})^3}{E_{var}^{3/2}}, \tag{2}$$

where n represents the data quantity, and E_i represents the electricity consumption of i.

Coefficient of kurtosis C_k reveals the existence of extreme values in the data list [45], calculated using Equation (3).

$$C_k = \frac{\frac{1}{n}\sum\limits_{i=1}^{n}(E_i - E_{mean})^4}{E_{var}^2} - 3, \tag{3}$$

where the parameters are the same as those in Equation (2).

Additionally, a simple dimensionless parameter, coefficient of the median C_m, is proposed in the present study, defined as the ratio of the difference between the maximum and median to the range, calculated using Equation (4). The theoretical and practical flow underlying the proposition of C_m is discussed in Section 4.2.

$$C_m = \frac{E_{max} - E_{median}}{E_{range}}. \tag{4}$$

The parameters mentioned above were calculated or recorded to understand the numerical characteristics of energy consumption for each device. Afterward, the parameters were compared to determine if their disparities could be used to directly recognize the variable-speed ME.

3. Results

3.1. ME Energy Consumption Distributions

The energy consumption distributions of the 16 ME items are shown in Figure 2. Some obvious differences could be witnessed among the devices. Taking the nine pumps as an example, some of the pumps consumed as much as >40 kWh electricity, such as GWP-3, while some consumed much less electricity, e.g., ChWP-1 and CWP-2. From the perspective of distribution, for some pumps, such as ChWP-3 and CWP-3, most of the energy consumption data were grouped near the maximal value. The high-energy-consumption data appeared hundreds of times more often than low-energy-consumption data did. In comparison, for other pumps, such as ChWP-1 and ChWP-2, energy consumption data were distributed more uniformly, while for GWP-1 and GWP-2, several different peaks could be seen in the distribution. The differences in energy consumption distribution among the ME items may provide inspiration for the recognition of variable-speed equipment. More detailed information on numerical analysis can be found in Section 3.2.
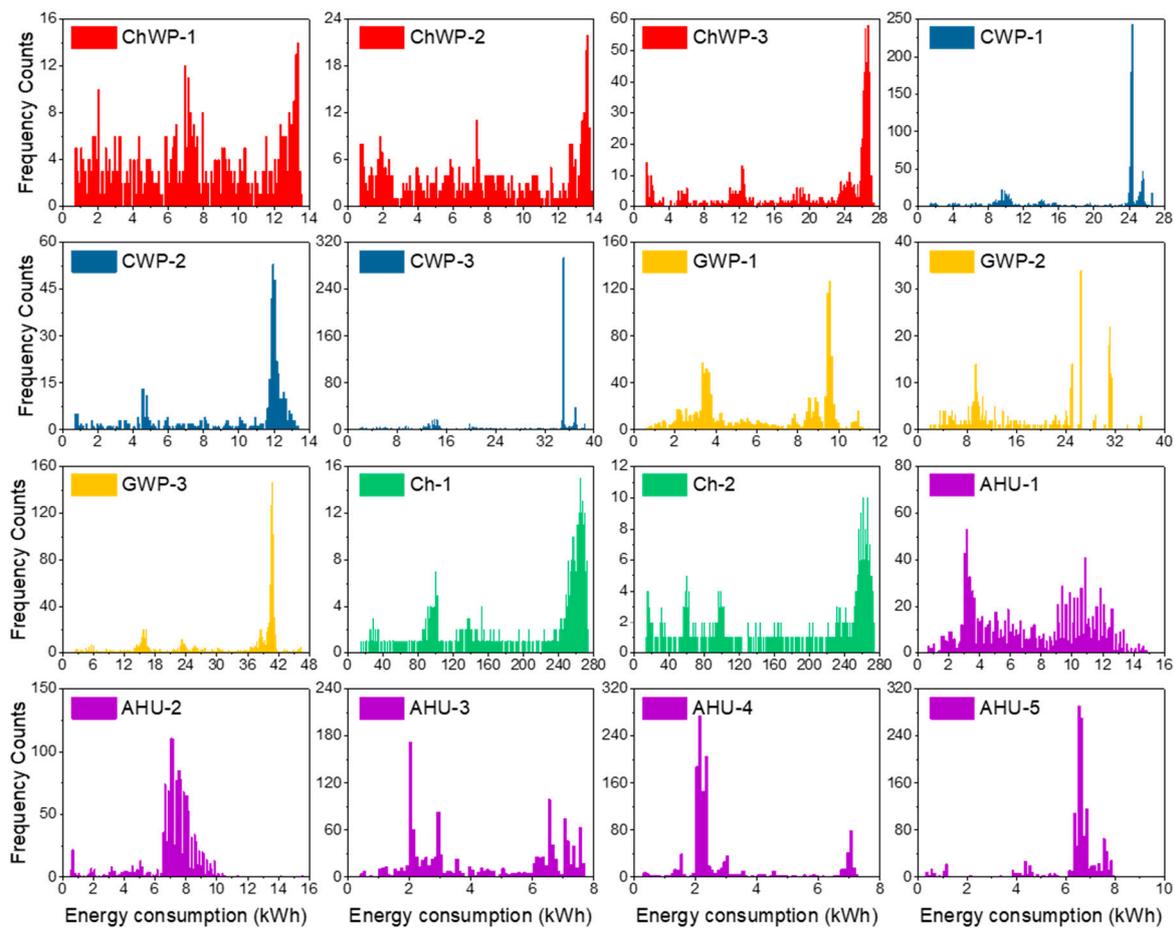
Figure 2. Energy consumption data distributions of 16 motor-driven equipment (ME) items in the case study.

3.2. Numerical Characteristics of Energy Consumption Data

The numerical parameters of the energy consumption data for each ME item were calculated, and they are listed in Table 1. As seen, most dimensional parameters of the AHUs were generally the lowest among all the ME, while the two chillers presented most of the highest dimensional parameters. Additionally, sorting the devices on the basis of these dimensional parameters resulted in individual differences. When referring to dimensionless parameters, the lowest C_s and C_v and the highest C_k were found for AHU-5, indicating that the energy consumption data were highly concentrated near the maximal value. In comparison, for ChWP-1, ChWP-2, GWP-1, GWP-2, AHU-1, and AHU-3, C_s and C_v were much higher, while C_k was much lower, implying more uniform distributions.

Table 1. Numerical characteristics of energy consumption distribution of 16 ME items in the case study. ChWP, chilled-water pump; CWP, condensing-water pump; GWP, glycol water pump; Ch, chiller; AHU, air-handling unit.

Equipment No.	E_{range}	E_{mean}	E_{sd}	E_{var}	C_s	C_k	E_{median}	E_{mode}	C_v	C_m
ChWP-1	12.80	7.45	3.93	15.47	−0.03	−1.21	7.20	6.90	0.53	0.49
ChWP-2	13.10	7.72	4.37	19.12	−0.07	−1.37	7.50	13.50	0.57	0.48
ChWP-3	25.90	19.97	8.23	67.80	−0.99	−0.45	24.60	26.30	0.41	0.10
CWP-1	25.20	19.69	7.14	50.91	−1.00	−0.55	24.20	24.30	0.36	0.10
CWP-2	12.60	9.70	3.58	12.84	−1.15	−0.13	11.80	12.00	0.37	0.12
CWP-3	36.50	28.12	10.64	113.23	−1.01	−0.55	35.00	35.00	0.38	0.10
GWP-1	10.80	6.37	3.02	9.15	−0.14	−1.61	6.70	9.50	0.47	0.44
GWP-2	34.40	18.26	9.56	91.41	0.03	−1.49	17.95	26.40	0.52	0.53
GWP-3	44.00	33.64	11.02	121.38	−1.25	0.16	40.30	40.70	0.33	0.14
Ch-1	258.00	216.08	71.66	5135.15	−1.19	−0.09	255.70	264.50	0.33	0.07
Ch-2	259.80	193.17	86.76	7527.56	−0.76	−1.09	249.60	258.50	0.45	0.09
AHU-1	14.10	7.74	3.58	12.80	−0.12	−1.35	8.50	10.80	0.46	0.45
AHU-2	14.90	7.11	1.78	3.17	−1.70	4.50	7.30	7.00	0.25	0.55
AHU-3	7.50	4.44	2.23	4.95	0.06	−1.65	3.90	2.00	0.50	0.52
AHU-4	7.30	2.97	1.73	2.99	1.60	1.15	2.20	2.20	0.58	0.74
AHU-5	8.80	6.26	1.47	2.17	−2.60	6.79	6.60	6.50	0.24	0.30

3.3. Recognition of Variable-Speed ME

3.3.1. Recognition of Variable-Speed Pumps

As seen in Table 1, the values of C_v and C_m allowed for dividing the pumps into two groups: one with comparatively high C_v and C_m, including ChWP-1, ChWP-2, GWP-1, and GWP-2, and another with comparatively low C_m and C_v, consisting of ChWP-3, CWP-1, CWP-2, CWP-3, and GWP-3. To highlight the two groups, a scatter matrix based on numerical parameters is shown in Figure 3.

Figure 3 clearly shows that the nine pumps were divided into two groups, as shown in row 9 and column 5. The four red points in the top-right corner represent ChWP-1, ChWP-2, GWP-1, and GWP-2. This high value indicated that the energy consumption data of the four pumps were not concentrated at the extrema but were distributed fairly uniformly, which was in accordance with the operation characteristics of variable-speed pumps. Therefore, it was assumed that ChWP-1, ChWP-2, GWP-1, and GWP-2 were variable-speed ME, while the five other pumps were constant-speed ME.

With the information transmitted from the field, the assumption was then validated, indicating the successful recognition of variable-speed pumps by this study.

C_m and C_s could be used to individually distinguish the two groups of pumps, as seen in row 4 combined with column 5 and row 9, respectively, while other numerical parameters failed. However, the calculation of C_m was much easier than that of C_s. Therefore, C_m may be a good parameter to be used as an indicator in the recognition of variable-speed pumps.

3.3.2. Verifiable Recognition of Variable-Speed AHUs and Constant-Speed Chillers

From the information transmitted on site, the five AHUs were operated with variable speed, while the two chillers worked with constant speed. Therefore, the recognition of AHUs and chillers could be further investigated to verify the applicability of the two indicators of C_m and C_s. The values of C_m and C_s of each of the seven devices are visually compared in Figure 4. Interestingly, C_m could be used to distinguish variable-speed AHUs and constant-speed chillers by using a single threshold (0.09–0.30), whereas C_s failed to facilitate recognition despite its applicability for pumps. The possible reason for this is discussed in Section 4.2.

The results presented thus far support the idea that proposed parameter C_m could be a good indicator in the recognition of variable-speed equipment.
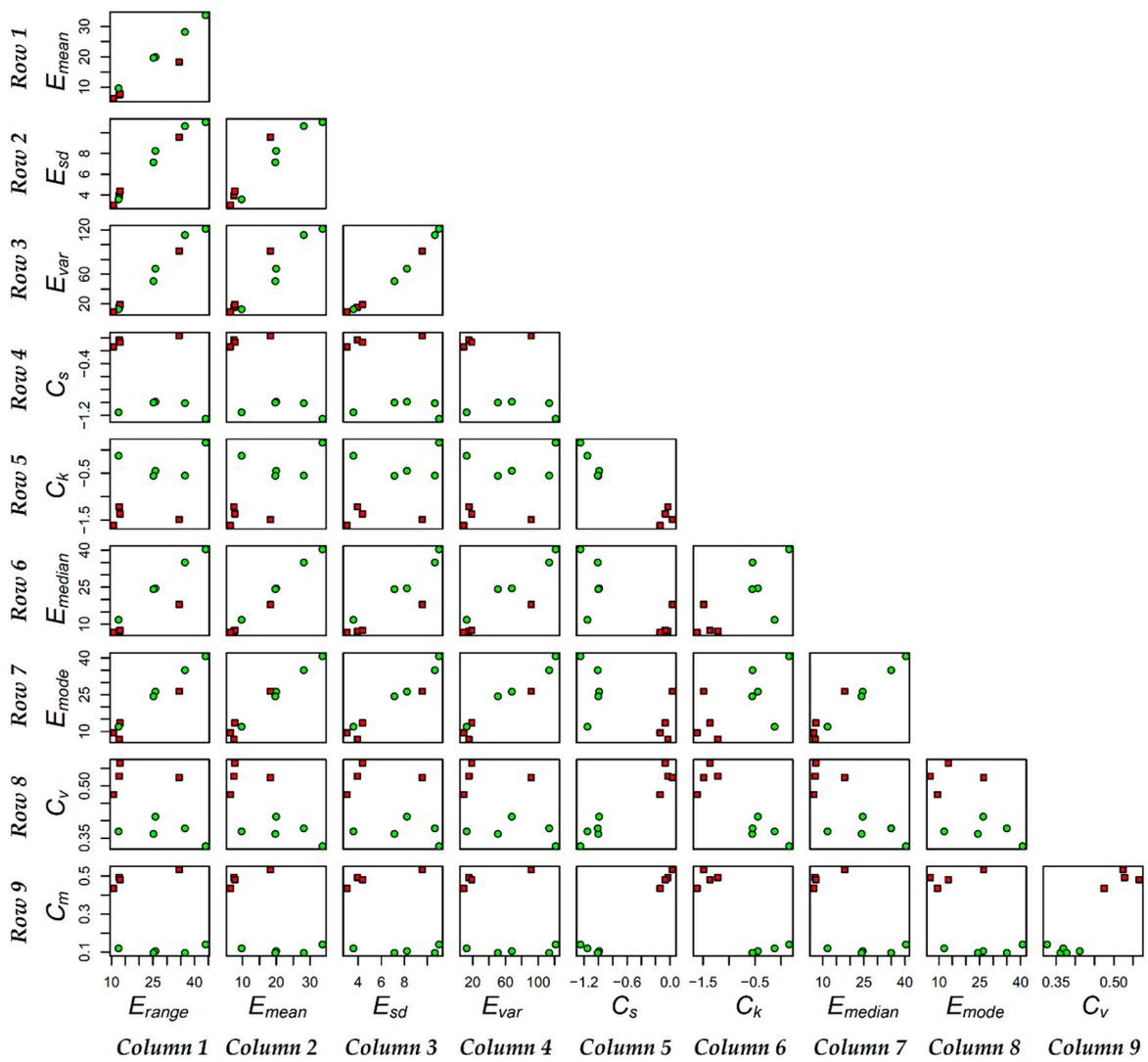
Figure 3. Scatter matrix based on numerical parameters of energy consumption data of nine pumps. Four red points in each scatter were recognized as variable-speed pumps, while five green points were recognized as constant-speed pumps.
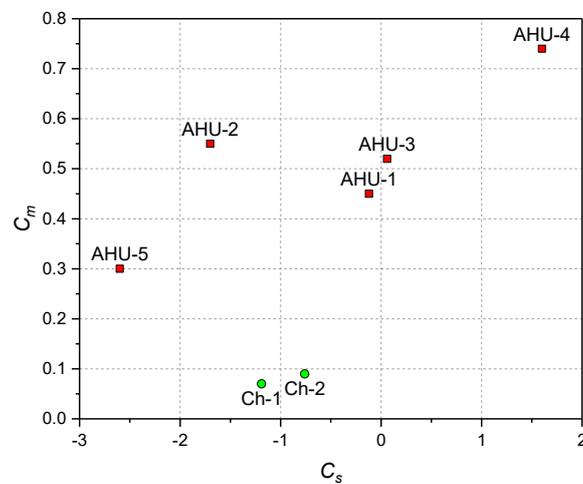
Figure 4. Values of C_m versus C_s for five AHUs and two chillers.

4. Discussions

4.1. Feature Selection of Numerical Characteristics of Energy Consumption Data

The above-mentioned results demonstrated that C_m, a numerical characteristic of the energy consumption data of an air-conditioning system, can be used to recognize whether pumps, chillers, and AHUs operate with variable speed. To verify and assess its effectiveness, feature selection was performed to compare the relevance of different numerical characteristics to determine whether the equipment worked with variable speed.

The feature selection method was originally based on a set of criteria or calculation methods to filter out feature fields for removal; then, the effectiveness of the remaining features was ranked relative to a specific target [46]. In this article, none of the 10 numerical characteristics mentioned above were removed. Instead, feature selection was used to find the most meaningful feature for the recognition and prediction of the target attribute from the 10 numerical characteristics of the energy consumption data. The feature selection considered one numerical parameter at a time to verify the performance of each parameter in recognizing and predicting the target attribute. The effectiveness value of each characteristic was calculated as $(1 - p)$, where p is the p-value of an appropriate statistical test of association between the candidate characteristic and the target attribute. In this investigation, all numerical characteristics were continuous variables, while the target attributes were categorical variables. Therefore, the numerical characteristics were set as observation variables, target attributes were set as control variables, and p-values based on the F-statistic were used. The idea was to perform a one-way analysis-of-variance (ANOVA) F-test for each characteristic to test whether the different categories of Y (representing the control variables) had the same means as the different categories of X (representing the observation variables). The p-value based on the F-statistic was calculated as p-value $= \text{Prob}\{F(J - 1, N - J) > F\}$, where $F(J - 1, N - J)$ is a random variable that follows F-distribution with degrees of freedom $J - 1$ and $N - J$, respectively; N is the total number of cases; J is the number of target-attribute categories, and

$$F = \frac{\sum\limits_{j=1}^{J} N_j(\bar{x}_j - \bar{\bar{x}})^2 / (J - 1)}{\sum\limits_{j=1}^{J} (N_j - 1)s_j^2 / (N - J)}, \tag{5}$$

where N_j represents the number of cases with $Y = j$; $\bar{x}_j$ represents the sample mean of characteristic X for target class $Y = j$; $\bar{\bar{x}}$ represents the grand mean of characteristic X; $\bar{\bar{x}} = \sum\limits_{j=1}^{J} N_j\bar{x}_j / N$, and s_j^2 represents the sample variance of characteristic X for target category $Y = j$. If the denominator in Equation (5) for a characteristic was zero, the p-value was set as 0 for that characteristic.

If the calculation result of $(1 - p)$ was larger than 0.95, denoting that the observation variable had strong correlation with the control variable, the tested numerical parameter was deemed effective for the recognition and prediction of the target attribute, and vice versa.

Table 2 lists the calculated ANOVA F-test p-values for the ten numerical characteristics ranked by their effectiveness relative to the attribute of variable-speed operations. The $(1 - p)$ values of C_m, E_{median}, E_{mean}, E_{mode}, E_{range}, and E_{sd} reached or exceeded the threshold for "effective" (0.95). However, E_{median}, E_{mean}, E_{mode}, E_{range}, and E_{sd} were dimensional characteristics; thus, the value strongly depended on the equipment and/or the system. For example, we could not infer or predict that equipment with different types, models, or specifications had the same target attribute only on the basis of the same values of E_{median}, E_{mean}, E_{mode}, E_{range}, or E_{sd} for energy consumption data. Therefore, the remaining dimensionless characteristic, C_m, was the only valid indicator that could be used for the recognition of variable- and constant-speed operation for different types of equipment in the present study.

Table 2. Feature selection of numerical characteristics of energy consumption data in terms of recognizing constant/variable-speed ME.

Rank	Numerical Characteristic	Effectiveness [1]	$(1 - p)$ Value
1	C_m	Effective	1.000
2	E_{median}	Effective	0.963
3	E_{mean}	Effective	0.959
4	E_{mode}	Effective	0.959
5	E_{range}	Effective	0.952
6	E_{sd}	Effective	0.950
7	E_{var}	Marginal	0.904
8	C_v	Ineffective	0.872
9	C_s	Ineffective	0.864
10	C_k	Ineffective	0.485

[1] Numerical characteristics labeled as effective, marginal, or ineffective on the basis of $(1 - p)$ value: $(1 - p) \geq 0.95$, effective; $0.95 > (1 - p) \geq 0.9$, marginal; $0.9 > (1 - p)$, ineffective.

With the above results and discussion, the effective applicability of proposed characteristic C_m was confirmed both practically and mathematically.

4.2. Construction Process of C_m

Initially, we also attempted to address the issue of recognition by using theoretical reasoning by implementing the intrinsic correlations of operational data. Variable-speed ME can better adjust to state changes than constant-speed ME can, thus making its operation more energy-efficient. However, for actual air-conditioning systems, it is not possible to simply use energy consumption data to determine that the comparatively low-energy-consumption ME is running at variable speed, or the comparatively high-energy-consumption ME is running at constant speed. Even if there are two identical devices used in parallel, it is possible that constant-speed ME may simply run for a shorter time than other ME through on–off control, resulting in lower energy consumption. Similarly, variable-speed ME may simply run longer, resulting in higher energy consumption. Furthermore, the energy consumption comparison across different brands, types, and models of ME is much more complex. It is also the case in real air-conditioning systems that the same type of ME consumes the same amount of energy, which obviously makes it impossible to distinguish between constant- and variable-speed operating states by using a comparison of their energy consumption. In addition, due to different resistance or loads in the system (or in the loop within the system), coupled with the fact that the operation of ME may be influenced by the different thermal, vibration, and electromagnetic environments, as well as by the power quality at the workplace, it is difficult to use energy consumption data to determine constant- or variable-speed ME via theoretical reasoning. Therefore, this study attempted to address this issue by investigating the distribution of energy consumption data.

Theoretically, there are at least three typical scenarios for the distribution of energy consumption data. The first scenario is symmetrical distribution, such as standard normal distribution that rarely occurs in real data. Instead, energy consumption data in real air-conditioning systems are either positively skewed (i.e., with a positive value of C_s) or negatively skewed (i.e., with a negative value of C_s).

Apparently, the energy consumption distributions of constant-speed operating equipment (i.e., ChWP-3, CWP-1, CWP-2, CWP-3, GWP-3, Ch-1, and Ch-2) in Figure 2 typically presented a negative skew; thus, their C_s values were negative, as listed in Table 1. However, it can also be seen from Table 1 that some variable-speed operating equipment (i.e., ChWP-1, ChWP-2, GWP-1, AHU-1, AHU-2, and AHU-5) had negative C_s values. Thus, C_s cannot easily and effectively distinguish variable- from constant-speed ME. C_s essentially measures the symmetry of distribution [47] or, more generally, characterizes the relationship between E_{mean} and E_{median} [48]. A positive C_s indicates that E_{mean} is larger than E_{median}, while a negative C_s implies the opposite. Additionally, a close-to-zero C_s reflects the

relative closeness of E_{mean} and E_{median}. Since all pumps were equipped on system-circulation loops, they tended to receive the same or similar linkage-control strategies/commands for operation to meet the system's flow-rate demands. Therefore, during the running period, the relationships between E_{mean} and E_{median} fell into two distinct groups according to variable- and constant-speed operations. This difference could be precisely captured by C_s. Nevertheless, other ME, especially terminals such as the AHUs in this study, are mainly regulated on the basis of the cooling demands of specific spaces. Compared with pumps, AHUs are generally not controlled in linkage; thus, the relationships between E_{mean} and E_{median} are intricate and difficult to characterize using C_s. This explains the disparity between the capability of C_s in distinguishing variable- from constant-speed operations for pumps and other ME (AHUs and chillers).

As demonstrated in Figure 2, the energy consumption of equipment working with a constant speed was generally fixed, and the distribution of energy consumption was thus concentrated near maximal value E_{max} over the running period. On the contrary, the distribution of energy consumption for variable-speed equipment was dispersed rather than focused around E_{max}. Therefore, the difference between maximal value E_{max} and the measure of central tendency, which was a simpler and more intuitive parameter, could compensate for the above shortcomings of C_s. Furthermore, with respect to the selection of the indicator of central tendency, among the three options of mode, E_{mode}, median E_{median}, and mean E_{mean}, actual energy consumption may have multiple modes, while the mean is sensitive to extremes (including outliers). Therefore, median E_{median}, a better measure for skewed data like energy consumption data, was chosen to eliminate these adverse effects. Then, considering the different numerical scales of energy consumption for different types, models, or specifications of equipment, the previous E_{max}–E_{median} was divided by the range E_{range} of energy consumption data to obtain dimensionless parameter C_m. The process described so far is exactly what was investigated at the theoretical and practical level when proposing the coefficient of the median C_m in the present study.

4.3. Scope of Application of C_m and Contributions of This Study

The contributions of this research and applicability of C_m are summarized as follows:

- The successful recognition of variable- and constant-speed ME using the proposed C_m in this paper addresses the research gap related to the reverse identification of equipment type through mining the energy consumption data of air-conditioning systems.
- The involved ME in this study had diverse components and brands and thus exhibited different distributions and scales of energy consumption data. Nevertheless, since differences in energy consumption scales were eliminated and made dimensionless in the development of parameter C_m, the method and results were not expected to be influenced by the specific brands, types, models, or power ratings of ME. Consequently, the proposed C_m is supposed to have versatility in general air-conditioning systems other than the specific example of this study.
- In addition, even if the constant- or variable-speed nature of the ME is already known, such an investigation could also diagnose if the variable-speed ME is working properly. In terms of diagnosis, by calculating the actual C_m value for a certain period of time and comparing it with the C_m value of the corresponding design conditions, it is possible to determine whether the operation of the ME is in accordance with the original design or to reflect on whether the design scheme is reasonable.
- The distinguishing results can provide essential information for the subsequent analysis of energy/cost saving potential by optimizing algorithms for different types of ME under their respective practical constraints. Moreover, the proposed C_m can also be of potential use to enhance the automatic processing of data, to reduce the direct involvement of field professionals, and to fundamentally support the automatic computer processing of massive amounts of air-conditioning system energy consumption data.

4.4. Limitations

However, this pilot study has the following limitations:

- The data involved in this study came from the case system's energy consumption monitoring platform. Therefore, there were no corresponding data available on the platform, such as the water flow rate of pumps, air volume of AHUs, cooling capacity of chillers, or related frequency data of inverters, which could be employed to verify the recognition results of the present case. In addition, energy consumption data were collected at an interval of 1 h, which seemed somewhat long for analysis in this paper, since constant-speed ME might accomplish several on/off cycles or the variable-speed ME might change its speed of rotation several times. This was due to the local preferential tariff for the ice-storage air-conditioning system, where the minimal interval between two adjustments for the ME was 1 h, including on/off control or changes in operating speed related to changes in energy consumption [49]. Therefore, the above limitations had negligible impact on our study, as the probed results demonstrated the effectiveness of the proposed indicator parameter in the recognition of variable-speed ME.

- The proposed coefficient of the median C_m could effectively recognize variable- and constant-speed operation modes for ME in the current study. Theoretically, the value of C_m ranges from 0 to 1. However, as stated in Section 4.2, the operation of real-world ME is influenced by the device and by the system to which it belongs, as well as by the thermal, vibration, and electromagnetic environments and the power quality of the environment in which it is located; hence, it is difficult to theoretically conclude a threshold value for recognition. On the basis of results to date, the confidence interval of the C_m value for ME running at constant speed ranged from 0.07 to 0.13 with 99% confidence, while the confidence interval of the C_m value for ME operating at variable speed ranged from 0.37 to 0.63 in this pilot study. Further study is warranted to collect more energy consumption data from different ME types and to theoretically investigate factors influencing C_m so as to give a more precise threshold value.

5. Conclusions

This study collected the energy consumption data of ME in an air-conditioning system in a commercial building. Numerical analysis was conducted to investigate the data characteristics of different ME types and to recognize variable-speed components. An indicator parameter, coefficient of the median C_m, was introduced. Results showed that the energy consumption data distributions of variable- and constant-speed ME were quite different. Coefficients of skewness C_s and of the median C_m could be used to distinguish variable- from constant-speed pumps. However, only proposed characteristic C_m could be used to successfully distinguish all different types of ME in terms of constant- and variable-speed operation. Moreover, the effective applicability of C_m was further mathematically confirmed by using a feature selection method.

Regardless of the small sample size of this study, the successful recognition of variable- and constant-speed ME using the proposed characteristic in this paper addresses the research gap related to the reverse identification of equipment type. The results are also of potential use for enhancing the automatic processing of data, reducing the direct involvement of field professionals, and supporting the foundation for the automatic computer processing of massive amounts of air-conditioning system energy consumption data.

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

References

1. Pérez-Lombard, L.; Ortiz, J.; Pout, C. A review on buildings energy consumption information. *Energy Build.* **2008**, *40*, 394–398. [CrossRef]
2. Afram, A.; Janabi-Sharifi, F. Theory and applications of HVAC control systems—A review of model predictive control (MPC). *Build. Environ.* **2014**, *72*, 343–355. [CrossRef]
3. Fong, K.F.; Hanby, V.I.; Chow, T.T. HVAC system optimization for energy management by evolutionary programming. *Energy Build.* **2006**, *38*, 220–231. [CrossRef]
4. Qureshi, T.Q.; Tassou, S.A. Variable-speed capacity control in refrigeration systems. *Appl. Eng.* **1996**, *16*, 103–113. [CrossRef]
5. Khalid, N. Efficient energy management: Is variable frequency drives the solution. *Procedia Soc. Behav. Sci.* **2014**, *145*, 371–376. [CrossRef]
6. Bose, B.K. Global energy scenario and impact of power electronics in 21st century. *IEEE Trans. Ind. Electron.* **2013**, *60*, 2638–2651. [CrossRef]
7. Bernier, M.A.; Bourret, B. Pumping energy and variable frequency drives. *Ashrae J.* **1999**, *41*, 37–40.
8. Zhou, Y.P.; Wu, J.Y.; Wang, R.Z.; Shiochi, S. Energy simulation in the variable refrigerant flow air-conditioning system under cooling conditions. *Energy Build.* **2007**, *39*, 212–220. [CrossRef]
9. Ma, Z.; Wang, S. Energy efficient control of variable speed pumps in complex building central air-conditioning systems. *Energy Build.* **2009**, *41*, 197–205. [CrossRef]
10. Roulet, C.A. Indoor environment quality in buildings and its impact on outdoor environment. *Energy Build.* **2001**, *33*, 183–191. [CrossRef]
11. Wang, S.; Jin, X. Model-based optimal control of VAV air-conditioning system using genetic algorithm. *Build. Environ.* **2000**, *35*, 471–487. [CrossRef]
12. Kusiak, A.; Li, M.; Tang, F. Modeling and optimization of HVAC energy consumption. *Appl. Energy* **2010**, *87*, 3092–3102. [CrossRef]
13. Lu, L.; Cai, W.; Soh, Y.C.; Xie, L.; Li, S. HVAC system optimization-Condenser water loop. *Energy Convers. Manag.* **2004**, *45*, 613–630. [CrossRef]
14. Shao, S.; Shi, W.; Li, X.; Chen, H. Performance representation of variable-speed compressor for inverter air conditioners based on experimental data. *Int. J. Refrig.* **2004**, *27*, 805–815. [CrossRef]
15. Kaya, D.; Yagmur, E.A.; Yigit, K.S.; Kilic, F.C.; Eren, A.S.; Celik, C. Energy efficiency in pumps. *Energy Convers. Manag.* **2008**, *49*, 1662–1673. [CrossRef]
16. Lönnberg, M. Variable Speed Drives for energy savings in hospitals. *World Pumps* **2007**, *2007*, 20–24. [CrossRef]
17. Montagud, C.; Corberán, J.M.; Montero, A. In situ optimization methodology for the water circulation pumps frequency of ground source heat pump systems. *Energy Build.* **2014**, *68*, 42–53. [CrossRef]
18. Teitel, M.; Levi, A.; Zhao, Y.; Barak, M.; Bar-lev, E.; Shmuel, D. Energy saving in agricultural buildings through fan motor control by variable frequency drives. *Energy Build.* **2008**, *40*, 953–960. [CrossRef]
19. Schibuola, L.; Scarpa, M.; Tambani, C. Variable speed drive (VSD) technology applied to HVAC systems for energy saving: An experimental investigation. *Energy Procedia* **2018**, *148*, 806–813. [CrossRef]
20. Yu, F.W.; Chan, K.T. Environmental performance and economic analysis of all-variable speed chiller systems with load-based speed control. *Appl. Eng.* **2009**, *29*, 1721–1729. [CrossRef]
21. Hartman, T. All-variable speed centrifugal chiller plants. *Ashrae J.* **2001**, *43*, 43–53.
22. Wang, F.J.; Chang, T.B.; Chen, M.T. Energy conservation for chiller plants by implementation of variable speed driven approach in an industrial building. *Energy Procedia* **2014**, *61*, 2537–2540. [CrossRef]
23. Del Col, D.; Azzolin, M.; Benassi, G.; Mantovan, M. Energy efficiency in a ground source heat pump with variable speed drives. *Energy Build.* **2015**, *91*, 105–114. [CrossRef]

24. Lee, C.K. Dynamic performance of ground-source heat pumps fitted with frequency inverters for part-load control. *Appl. Energy* **2010**, *87*, 3507–3513. [CrossRef]

25. Yu, Z.; Haghighat, F.; Fung, B.C.M.; Zhou, L. A novel methodology for knowledge discovery through mining associations between building operational data. *Energy Build.* **2012**, *47*, 430–440. [CrossRef]

26. Ahmed, A.; Korres, N.E.; Ploennigs, J.; Elhadi, H.; Menzel, K. Mining building performance data for energy-efficient operation. *Adv. Eng. Inform.* **2011**, *25*, 341–354. [CrossRef]

27. Kim, H.; Stumpf, A.; Kim, W. Analysis of an energy efficient building design through data mining approach. *Autom. Constr.* **2011**, *20*, 37–43. [CrossRef]

28. 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]

29. Ma, R.; Yu, N. A New Route for Energy Efficiency Diagnosis and Potential Analysis of Energy Consumption from Air-Conditioning System. In Proceedings of the 14th International Conference for Enhanced Building Operations (ICEBO2014), Beijing, China, 14–17 September 2014.

30. Du, Z.; Fan, B.; Jin, X.; Chi, J. Fault detection and diagnosis for buildings and HVAC systems using combined neural networks and subtractive clustering analysis. *Build. Environ.* **2014**, *73*, 1–11. [CrossRef]

31. Du, Z.; Jin, X.; Yang, Y. Fault diagnosis for temperature, flow rate and pressure sensors in VAV systems using wavelet neural network. *Appl. Energy* **2009**, *86*, 1624–1631. [CrossRef]

32. Lee, W.-Y.; House, J.M.; Kyong, N.-H. Subsystem level fault diagnosis of a building's air-handling unit using general regression neural networks. *Appl. Energy* **2004**, *77*, 153–170. [CrossRef]

33. Chen, Y.; Hao, X.; Zhang, G.; Wang, S. Flow meter fault isolation in building central chilling systems using wavelet analysis. *Energy Convers. Manag.* **2006**, *47*, 1700–1710. [CrossRef]

34. Yang, X.; Du, Z.; Jin, X.; Fan, B.; Guo, Y.; Yang, Y. Wavelet Packet and Neural Network Methods for Fixed Bias Fault Diagnosis for Air Handling Unit Sensors. In Proceedings of the 6th International Symposium on Heating, Ventilating and Air Conditioning, ISHVAC 2009, Nanjing, China, 6–9 November 2009.

35. Wang, S.; Cui, J. Sensor-fault detection, diagnosis and estimation for centrifugal chiller systems using principal-component analysis method. *Appl. Energy* **2005**, *82*, 197–213. [CrossRef]

36. Jin, X.; Du, Z. Fault tolerant control of outdoor air and AHU supply air temperature in VAV air conditioning systems using PCA method. *Appl. Eng.* **2006**, *26*, 1226–1237. [CrossRef]

37. Du, Z.; Jin, X.; Yang, Y. Wavelet Neural Network-Based Fault Diagnosis in Air-Handling Units. *HvacR Res.* **2008**, *14*, 959–973. [CrossRef]

38. Zhu, Y.; Jin, X.; Du, Z. Fault diagnosis for sensors in air handling unit based on neural network pre-processed by wavelet and fractal. *Energy Build.* **2012**, *44*, 7–16. [CrossRef]

39. Idowu, S.; Saguna, S.; Ahlund, C. Applied machine learning: Forecasting heat load in district heating system. *Energy Build.* **2016**, *133*, 478–488. [CrossRef]

40. Zhao, D.; Zhong, M.; Zhang, X.; Su, X. Energy consumption predicting model of VRV (Variable refrigerant volume) system in office buildings based on data mining. *Energy* **2016**, *102*, 660–668. [CrossRef]

41. Liu, J.; Wang, J.; Li, G.; Chen, H.; Shen, L.; Xing, L. Evaluation of the energy performance of variable refrigerant flow systems using dynamic energy benchmarks based on data mining techniques. *Appl. Energy* **2017**, *208*, 522–539. [CrossRef]

42. 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]

43. Brown, C.E. *Applied Multivariate Statistics in Geohydrology and Related Sciences*; Springer: Berlin/Heidelberg, Germany, 1998; ISBN 978-3-642-80330-7.

44. Groeneveld, R.A.; Meeden, G. Measuring Skewness and Kurtosis. *Statistics* **1984**, *33*, 391–399. [CrossRef]

45. Joanes, D.N.; Gill, C.A. Comparing measures of sample skewness and kurtosis. *J. R. Stat. Soc. Ser. D Stat.* **1998**, *47*, 183–189. [CrossRef]

46. Guyon, I.; Elisseeff, A. An introduction to variable and feature selection. *J. Mach. Learn. Res.* **2003**, *3*, 1157–1182.

47. Newton, R.R.; Rudestam, K.E. *Your Statistical Consultant: Answers to Your Data Analysis Questions*; SAGE Publications, Ltd.: London, UK, 2013; ISBN 9781412997591.

48. Lewis-Beck, M.; Bryman, A.; Futing Liao, T. *The SAGE Encyclopedia of Social Science Research Methods*; Sage Publications, Inc.: Thousand Oaks, CA, USA, 2004; ISBN 9780761923633.
49. 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.