Energy and Buildings 139 (2017) 653–664

Contents lists available at ScienceDirect

Energy and Buildings

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

Critique of operating variables importance on chiller energy performance using random forest

F.W. Yu [a,*], W.T. Ho [a], K.T. Chan [b], R.K.Y. Sit [a]

[a] Hong Kong Community College, The Hong Kong Polytechnic University, 8 Hung Lok Road, Hung Hom, Hong Kong
[b] Department of Building Services Engineering, The Hong Kong Polytechnic University, Hung Hom, Hong Kong

ARTICLE INFO

Article history:
Received 28 September 2016
Received in revised form 1 January 2017
Accepted 19 January 2017
Available online 22 January 2017

Keywords:
Air-cooled chiller
Coefficient of performance
Heat rejection
Random forest

ABSTRACT

Chiller systems take up the major proportion of electricity used in commercial buildings. Their energy performance in terms of coefficient of performance (COP) depends on how the controllable and uncontrollable variables change. The aim of this study is to use the random forest (RF) method to measure variables importance and predict the COP. A sophisticated data trend log was carried out on an air-cooled chiller with advanced heat rejection features. The variables measured are: the flow rate of chilled water; the supply and return temperatures of chilled water; the temperature and relative humidity of outdoor air; the compressor power; the evaporating temperature; the condensing temperature; the number and speed of condenser fans staged; the temperatures of air entering and leaving the condenser. The data were logged at 5-min intervals in Aug 2015–Mar 2016. The RF models for different operating modes were validated, with a robust coefficient of determination of 80.52–96.53% for the testing data set. The chiller part load ratio, the condensing temperature, the chilled water flow rate, the heat rejection airflow rate and the wet-bulb temperature are the top five important variables in the prediction of COP. Yet they are not fully considered in typical regression models. Results of this study provide an insight into which variables are important to predict accurately the COP under different energy efficient features. The need of identifying the changing pattern of important variables is ascertained.

1. Introduction

Providing cooling energy by air-conditioning systems is energy intensive. As electricity is one of the most reliable and convenient power sources, chillers working under the vapour compression cycle are commonly used to convert electricity into cooling energy. As mentioned by Huang et al. [1], Yu and Chan [2] and Tran et al. [3], the operation of chiller systems in commercial buildings could take up one-third to half of the total electricity consumption. The energy performance of chillers is usually described by the coefficient of performance (COP) which is the cooling energy output in kW divided by the electric power input in kW. Operating chillers with the maximum COP as far as possible is the key to minimizing electricity consumption.

As Lee et al. [4] mentioned, simulation is often used to predict and optimize the COP with respect to a set of operating variables. Robust design of chiller systems includes a central control and monitoring system which facilitates a sophisticated trend log of operating data for performance analysis. Grey and black-box modelling techniques are considered when massive data are available for operating variables with complicated interaction and under the non-linear processes. According to Labus et al. [5], Lee et al. [6] and Wei and Xu [7], multivariate linear regression (MLR) and artificial neural network (ANN) are common grey and black-box modelling methods used for chiller system simulation. Accuracy is the primary concern on selecting the modelling methods, followed by the identification of variables importance, the ease of application and time of computation, etc. Depending on the nature of operating variables, various algorithms embedded in the modelling methods may be considered to reach optimal results of the operating variables. Regarding ANN models, the relationship between the input and output variables are hidden and cannot be identified explicitly. While the MLR models can illustrate the correlation between the input and output variables by regression coefficients, the accuracy of MLR models depends on which input variables and their multivariate terms (e.g. outdoor temperature x outdoor temperature, outdoor temperature x relative humidity, outdoor temperature x chilled water flowrate, etc.) are included. The degree of input variables may also affect the accuracy. A dedicated effort is required to

* Corresponding author.
E-mail address: ccyufw@hkcc-polyu.edu.hk (F.W. Yu).

http://dx.doi.org/10.1016/j.enbuild.2017.01.063
0378-7788/

Nomenclature

COP	Coefficient of performance
c_{pa}	Specific heat capacity of air (1.02 kJ/kg °C)
c_{pw}	Specific heat capacity of water (4.19 kJ/kg °C)
E_{cc}	Compressor power (kW)
E_{cf}	Condenser fan power (kW)
E_{mist}	Mist pump power (kW)
E_{total}	Chiller total electric power (kW)
F_{cf}	Power frequency of variable speed drive (Hz)
MSE	Mean square error of the output variable
m_w	Chilled water mass flow rate (kg/s)
N_{cf}	Number of condenser fans operating
Q_{cd}	Heat rejection (kW)
Q_{cl}	Cooling capacity (kW)
$Q_{cl,nom}$	Nominal cooling capacity (282 kW)
PLR	Part load ratio
RH	Relative humidity (%)
T_{cd}	Saturated condensing temperature (°C)
T_{cdae}	Temperature of air entering the condenser (°C)
T_{cdal}	Temperature of air leaving the condenser (°C)
T_{chwr}	Temperature of return chilled water (°C)
T_{chws}	Temperature of supply chilled water (°C)
T_{ev}	Saturated evaporating temperature (°C)
T_o	Dry bulb temperature of outdoor air (°C)
T_w	Wet bulb temperature of outdoor air (°C)
V_a	Heat rejection airflow rate (m^3/s)
ρ_a	Air density (1.2 kg/m^3)
c_{pa}	Specific heat capacity of air (1.02 kg/kJ °C)
η_{mist}	Cooling effectiveness of mist system

examine the optimum order and combination of input variables to predict the outputs with the highest accuracy.

The random forest (RF) models developed by Liaw and Wiener [8] make use of classifying input variables by decision tree methods to identify variables importance and predict output variables with high accuracy. The RF algorithm can handle a huge set of input variables subject to certain outliers and noise in data. Goudarzi et al. [9] indicated that RF models were found to have superior accuracy compared with the MLR and ANN ones in predicting the retention indices of some polycyclic aromatic hydrocarbons. In the study of building age prediction and building energy modelling by Tooke et al. [10], RF regression provided a substantial advantage over more typical classification routines when predicting the building age, due to an unbiased surrogate for cross-validation with out-of-bag data. Tsanas and Xifara [11] concluded that the RF method outperformed the iteratively reweighted least squares method in evaluating an accurate functional relationship between the input and output variables. Li et al. [12] considered using RF with their tree-structured learning model to capture a more enhanced structure of common faults for the fault detection and diagnosis of building cooling systems. Yan et al. [13] ascertained the strength of decision tree classification with RF to develop a data-driven diagnostic strategy for air-handling units and to interpret fault features.

The aim of this study is to use the RF method to analyze the importance of operating variables on the COP of an air-cooled chiller and the COP prediction with high accuracy. The experimental set up of the chiller will be described. The RF method and the model development will be presented. Experimental results and the RF model will be discussed. Correlations among the variables will be investigated. The significance of this study is to demonstrate how an RF model facilitates the identification of important variables to improve the accuracy of the COP prediction. The technique is generic and applicable to other types of building engineering systems.

2. Method of study

2.1. Description of the chiller

The chiller was a lead chiller operating with three large chillers of the same type in an institutional building. It used the refrigerant R134a and its nominal capacity and full load COP are 282 kW and 2.8, respectively. Fig. 1 shows a schematic diagram of the chiller created by the authors. The diagram shows the physical arrangement of components: the evaporator, the two compressors, the two expansion valves and the air-cooled condenser. It also contains a mist system to precool condenser air. The two refrigeration circuits were identical and the screw compressor in each circuit had three capacity steps controlled by regulating the sliding valves. Refrigerant control was done by using electronic expansion valves to maintain a degree of superheat of 5 °C at the exit of the evaporator. Each circuit contained three identical condenser fans, each of which had a rated power of 2.4 kW and delivered a heat rejection airflow rate of 3.28 m^3/s at full speed. A common set of chilled water pumps was used to deliver chilled water to each operating chiller. Depending on the combination of the operating chillers and pumps, a varying flow rate was identified at the chiller. To inves-

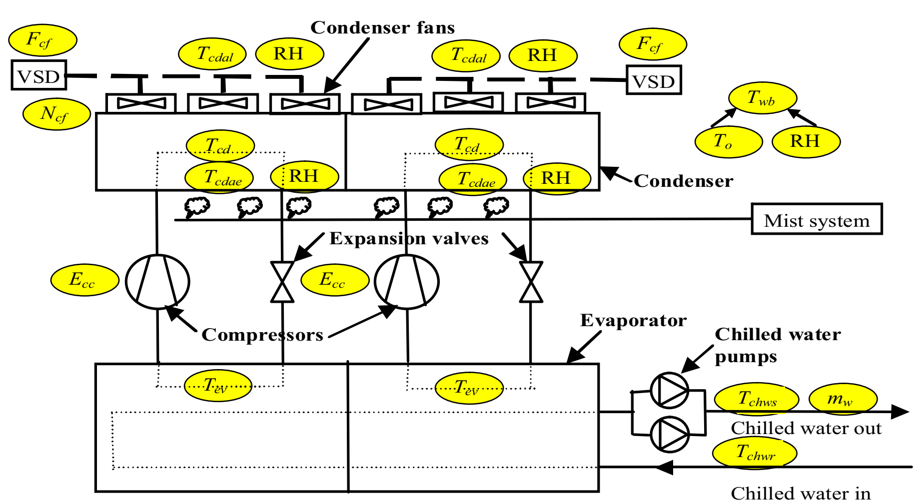

Fig. 1. Schematic diagram of the chiller with measurement locations of operating variables.