ETM0011

Data Fusion Technique for Leveraging Reliability of Supermarket Field Data

D. Woradechjumroen[1*], Taperit Tongshoob[1]

[1] Department of Mechanical System Engineering and Industrial Innovation, School of Engineering, Sripatum University, Bangkok, Thailand
[*] Corresponding Author: denchai.wora@gmail.com, denchai.wo@spu.ac.th (+66)948503366

Abstract

Energy savings utilizing both energy management strategy and enabling technologies are significantly challenged by many researchers. Especially, supermarkets are one of the largest commercial sectors, in which at least 50% of the total power usages are consumed by heating, ventilation and air-conditioning and refrigeration (HVAC&R) systems. Collecting field test data from supermarkets has been efficiently used for the system performance analysis and evaluation because degradation or sudden faults could occur continuously on the HVAC&R in stores; the faults causes excessive power consumptions and decreased system efficiency if the machines are lacked of well-experienced maintenances and commissioning. However, the obtained field data may include faulty operation data, unsteady-state data, sensor measurement errors and uncertainties; all mentioned problems lead to inefficient analysis and unpredictable improvement plans. To this end, this paper presents a data fusion technique for reducing the data errors to improve data reliability for effective HVAC&R performance analysis. Two supermarket data consisting of acceptable and faulty conditions (store A and B) are tested by the proposed strategy. The technique is composed of the three stages. At first, the data are analyzed based the control functions of HVAC&R systems and driving force condition ranges (zone temperature, indoor relative humidity and outdoor temperature). Then, outlier identification based on z-scores (standard scores) is used to determine outliers from a normal power consumption. Thirdly, when an outlier is detected, the energy interaction ranges at the period of the identified abnormal variation are rechecked by the same approach to assure that the detection is not from the severe variation of outdoor temperature from low to very high value. Also, fixed indoor relative humidity ranges are used to reduce the effect to refrigeration systems. The proposed approach shows that the outliers can be detected frequently in store B since faulty operations and commissioning occur continuously due to unsuitable routine operations resulting in excessive power consumptions. The technique can be further used as one of the procedures to clean abnormal data conditions for fault detection and diagnosis (FDD) process in related areas.

Keywords: Data Fusion, Energy Interaction, Faulty Operations, HVAC&R, Outlier

1. Introduction

Supermarkets are one of the most electricity-intensive categories in commercial buildings. For instance in USA, they account approximately 7% of the energy use for the primary commercial sector (18.35 Quads) [1]. Specifically, 50% of each supermarket energy use is consumed by refrigeration systems and 20% is accounted by heating, ventilation and air-conditioning (HVAC) operations. These coupled system called HVAC&R often do not perform as well as design conditions because abrupt or degradation faults occur during initial installation and commissioning or they are developing during routine operation. With these happenings, they can incur waste energy consumption about 30% on average. With the engineering point of view, faults happening in this building type lead to challenges for developing enabling technologies using non-invasive techniques which do not interrupt original system performance or do not need to shot down the system for the monitoring or installation.

To be more specific, supermarkets are one of the most integrated complex systems since they include HVAC&R systems; these two systems are coupled in terms of energy interaction and are the integrated largest contributor of energy usage and electricity peak demand. The driving force conditions of refrigeration systems are designed at 95°F (35° C) for an outdoor condition and 75°F (24° C) and 55% RH for an indoor condition. Not only is outdoor air temperature (OAT) driving force influencing on the power consumption of refrigeration systems, but the indoor conditions, that are provided by HVAC systems such as zone air temperature (ZAT) and indoor air relative humidity (IARH), also affect the power consumption; refrigeration systems cannot be operated optimally without a well-conditioned space served by HVAC operations. The suitable indoor condition of refrigeration systems can happen when the sizing and suitable operations of HVAC systems are proper. The indoor conditions of the refrigeration systems provided by HVAC systems are called “energy interaction” [2]. Faulty interactions indirectly cause faulty operations in each subsystem. For example, the interactions between improper indoor conditions and refrigeration systems lead to excessive power consumption in supermarkets. For instance, HVAC systems provide inappropriate indoor condition in a supermarket (e.g., 60% RH and 78 °F) and then the refrigeration equipment (e.g., refrigeration compressors) can consume more energy because latent load is increased due to more IARH. Conversely, whenever IARH could be controlled between 55% and 35% RH at 75°F for a zone temperature, the power consumption of refrigeration systems can be reduced [3].

To minimize these improper routine operations of the systems, automated fault detection and diagnostic (AFDD) can automate the process of continuous commissioning and endow the building energy systems with intelligence; so that they can self-diagnose problems and even self-execute correcting actions for non-optimal operations or provide recommendation reports for building operators or service contractors to fix the problems promptly and optimally schedule preventative maintenances. However, efficient AFDD algorithms can be developed using fault-free data which can be obtained from the well-monitored sensors of lab testing. Supermarket lab is cost-prohibitive in practice. To replace this limitation for developing AFDD for system interaction in supermarkets, field test data with the well-designed and well-operated systems can be utilized via a system monitoring. Unfortunately, obtained field data could include faulty operations, unsteady-state conditions, sensor measurement errors and uncertainties; all aforementioned issues lead to unqualified and unsatisfied data for energy interaction model used for AFDD development.

To this end, this paper applies an efficient and easy-to-use process called "data fusion" which is one of smart building energy solutions technologies (S-BEST) procedures [4]; it leverages the data quality and reliability of a monitored system. In the paper, it firstly introduces related backgrounds for understanding the research contributions to further develop ongoing projects in near future. Then, standard score (z-score) is briefly explained for a proposed methodology of the research paper. In terms of the method implementation, two supermarkets having similar HVAC&R system operations are used as the case studies for developing improved energy interaction analysis. The proposed method will be used as a tool to further develop efficient AFDD algorithms for automatically analyzing faulty interactions in Thailand supermarkets.

2. Backgrounds

To leverage data reliability and reduce errors, this section briefly introduces Smart Building Energy Solutions Technologies (S-BEST) and energy interactions of supermarket equipment. According to the vision illustration as depicted in Fig. 1, S-BEST is the integrated modern energy management technologies for constructing network connections of energy systems, building environment, community and manufacturers through data exchange carrier or big data using web-based or on-line monitoring system (e.g. chiller monitoring systems). The procedures of S-BEST for leveraging field data quality are in section 2.1 and 2.2 as follows:

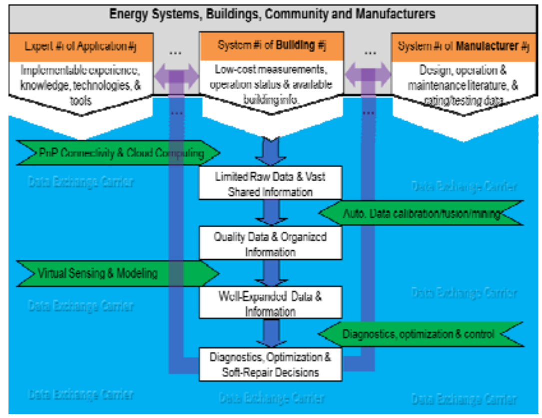

Fig. 1 Vision of S-BEST [5]

2.1 Data Exchange Carrier

The main objective of data exchange carrier is to construct data patterns of building energy systems through big data of the systems via web-based monitoring or cloud-based data storage. The building community connections can reduce the cost of use in terms of unnecessary data storage, can enhance more field-proven demonstration and are more scalable for optimally managing energy efficiency in new buildings and building renovation. The verified data patterns are very useful for further analyzing similar system operations.

Applying the data carrier to the current research, the main server of Smart Building Laboratory at University of Nebraska - Lincoln is connected to BACnet protocol of six supermarkets. The data of HVAC&R were exported from the server in comma separated value (CSV) format. To manage this big data efficiently, visual basic for applications (VBA) macro codes are created to sort and organize the data categories.

2.2 Automated calibration and data fusion

After obtaining the big data from the server, this is an automated process to permanently reduce or eliminate inherent errors occurring in physical sensors which are practically caused by bad location of a sensor, out-of-calibration and sensor failure. With the potential implementation of automated calibration [6], it can eliminate location-related errors, sensor-inherent errors, avoid utilizing failed measurements and replace failed measurements with alternatives. These avoidances lead to energy savings at field control levels of a building automation system (BAS) because sensor errors will result in waste energy consumption in buildings. For instance, a discharge air temperature sensor is wrongly read at 55 ºF (12.8 ºC); however, the true value is 53 ºF (11.7 ºC). This two degrees Fahrenheit of extra cooling translates to 5.8 kW of demand based on the operations of a fairly efficient chiller plant (0.8 system kW/ton) and a 40,000 cfm constant volume reheat air handler [7]; the extra cooling leads to 15,000 kWh per year or approximately equals to $1200 per year based on a typical 2,600 hour on the routine operation of an office building. In terms of data fusion, it is homogenous to data filter, so it is a useful tool to treat data quality of field tests used for developing and validating the fault-free models of model-based FDD algorithms.

2.3 Typical Energy Interaction

The supermarket refrigeration systems are designed based on the two driving conditions consisting of: 1) the indoor conditions (zone temperature and indoor RH) provided by HVAC systems and 2) outdoor air temperature. The faulty indoor conditions significantly affect the degraded efficiency and performance of refrigeration systems in supermarket leading to excessive power consumptions.

To investigate field test data, an energy interaction is a potential tool for preliminarily verifying data performance obtained from sensors [2, 5]. Typical interactions of the supermarket operations were proposed in [2]; however, it cannot diagnose simultaneous faults. The example of the typical interactions between considered parameters and equipment is explained as follow:

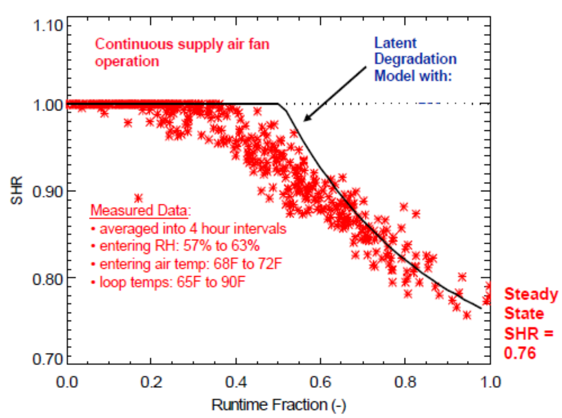

Fig. 2 Correlation between SHR and compressor runtime fraction [8]

Rooftop Unit (RTU) versus OAT, IARH and ZAT

The interaction between a RTU and IARH is not significant because indoor relative humidity is not the controlled parameter; the RTU is controlled by set points of temperature. However, IARH may strongly affect RTU compressor power consumption indirectly in case of inherent over-sizing issue. As a result, the latent capacity is not properly absorbed leading to high IARH. Fig. 2 illustrates the latent capacity degradation model proposed by Henderson [8]. This model is fitted by the correlation between the runtime fraction (RTF) of a compressor (see RTF definition in [9]) and the ratio of sensible capacity over total capacity (SHR). It can be noticed that RTF being 40% corresponds to SHR 1; it implies the latent load cannot be absorbed by the cooling coil of a RTU because the coil is saturated.

Applying this model, high IARH may be caused by the operations of oversized RTUs leading to higher power consumption in supermarkets.

The results of correlation investigation are tabulated in Table 1. The table concludes the R values between equipment power consumption and the three independent parameters [2].

Table 1 Expected R values of interactions

Equipment	Independent parameters		
	OAT	IARH	ZAT
RTU	Medium	Low	Medium
DHU	Medium	Medium	Low
Refrigeration	Medium	Medium	Medium
ASH	Low	High	Low

Note: low (0-0.50), medium (0.51-0.89) and high R value (0.9-1.0)

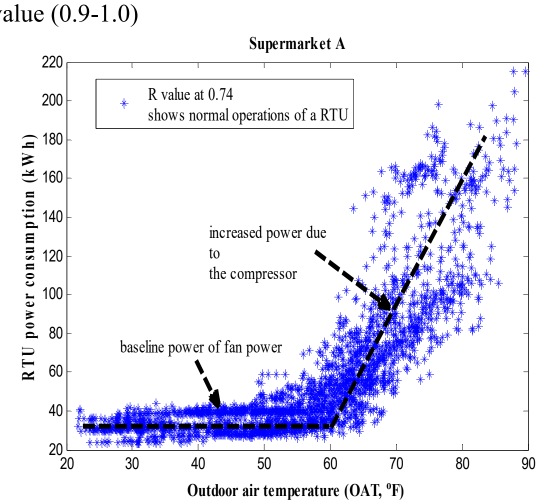

Fig.3 The plot between OAT and RTU in supermarket A

Fault Analysis Example (RTU vs. OAT and ZAT)

From Table 1, it is evident that the expected R value between IARH and RTU should be low because RTU controllers are typically functioned by a temperature set-point or ZAT. Meanwhile, a medium level of R value is expected between RTU and OAT or ZAT because these two parameters are the driving force conditions of a RTU.

For OAT and RTU, RTU power consumption practically increases as OAT is high, but the fluctuation could occur due to faulty sensors (e.g. OAT sensor) or controller malfunction resulting in lower efficiency. Fig. 3 shows the normal trend example of RTU operations consisting of: 1) baseline powers of RTU fans at about 60 °F (15.5 °C) and 2) increased power range when compressors are staged-on.

ETM0011

In contrast to Fig. 3, Fig. 4 depicts the skew data comparing to the normal trend. The area of the skew data could be caused by faults in condensers and/or compressors or even damper stuck faults. To isolate simultaneous faults, other measuring data such as refrigerant-side parameters are required to be monitored more in the field test.

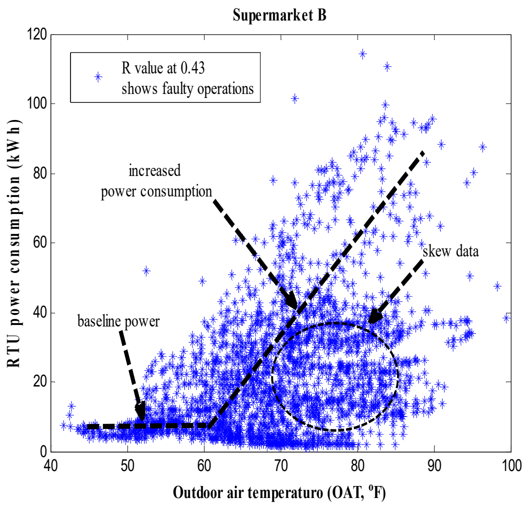

Fig. 4 The plot between OAT and RTU in supermarket B

3. Method

Data frequency depends on the available hard drive space of each BAS. In case of data limitation, the data frequency can be stored in short intervals around 1 – 5 minutes but for a limited amount time, 1 – 3 days. To keep the data longer, the sampling interval around 15 minutes is applied for longer amounts of time around 1-2 weeks. If the data are stored at short intervals for a limited amount of time, a user cannot see trends because the data will have already been replaced with new data. Additionally, if data are stored in large intervals, there can be a loss of granularity that is critical for analysis.

3.1 Data Sampling

For the analysis performed in this study, data were collected from each BAS at an average frequency of two minutes for the HVAC and refrigeration systems. Furthermore, the power meter data were collected at either one minute or 15 minute intervals. The data were stored on servers, with much larger capacities than a BAS, and no data were deleted from the servers. In order to reduce the amount of data to a manageable level, averages were taken from each sensor from each hour. Calculating the averages for each hour to drastically decrease the time, it took to perform the analysis and still maintained data integrity. Additionally, by averaging the power meter data, kilowatt hour (kWh) values were obtained instead of kilowatt values. The data including power consumptions of equipment, OAT, IARH and ZAT were obtained from 3/3/2011 to 11/30/2011 as tabulated in Table 2.

Table 2 Information of measured data obtained from BAS

Supermarket	IARH Range (%)	OAT Range(°F)	ZAT Range(°F)
A	20 - 63	50 - 90	61-72
B	25 - 70	42 - 99	67-75

3.2 Z-Score

From data sorting of the obtained CSV files in last section, six steps depicted in Fig. 5 are designed as follows:

Step 1 is to compute an average value ($\overline{x}$) and standard deviation (sd) of equipment kWh; the control equipment is staged on when OAT is more than 60 °F in each store for further quantifying abnormal kWh in terms of outliers.

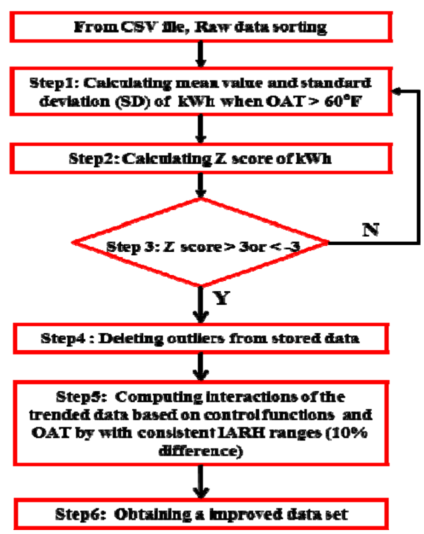

Fig. 5 Outlier Identification Procedure

Step 2 is further used to compute z-score in Eq. 1.

$$z = \frac{x_i - \overline{x}}{sd} \qquad (1)$$

, where $x_i = kWh_i$ of each machine at an identified sampling time

Step 3 is to identify outliers when computed z is more than the selected value, z > 3 or z < -3; if not, the process goes back to step 1 for rechecking the statistical values and re-selecting a new threshold.

Step 4 is utilized to reduce data errors by deleting outliers from the original data since most of excessive power consumptions in terms of the outliers are caused by faulty operations.

Step 5 is the verification process by examining that the identified periods of outliers are not influenced by severe OAT changes because high OAT may result in high power consumptions. The interaction between OAT and kWh is conducted for the verification by using consistent IARH difference range (e.g. 30-40%)

Step 6 is to improve data quality; three fixed IARH ranges are used to isolate IARH effect from refrigeration power consumption. With the specific data range, OAT versus kWh is assumed to be linear relation to examine equipment operations based on the control functions.

4. Method Implementation

For the method implementation, the two supermarkets – A and B with acceptable and faulty HVAC&R operations in Table 2, respectively, are used as the cases studies.

4.1 Data Fusion based OAT range

According to Fig. 3, the RTU compressors are staged on when OAT is more than 60 °F in store A which is similar to store B in Fig. 4. However, there are malfunctions in the RTU operations of store B, thus some skew data are deviated from the increased power consumption trend. First of all, OAT being more than 60 °F is used to consider the RTU operations and to reduce some outliers from the raw data in both stores.

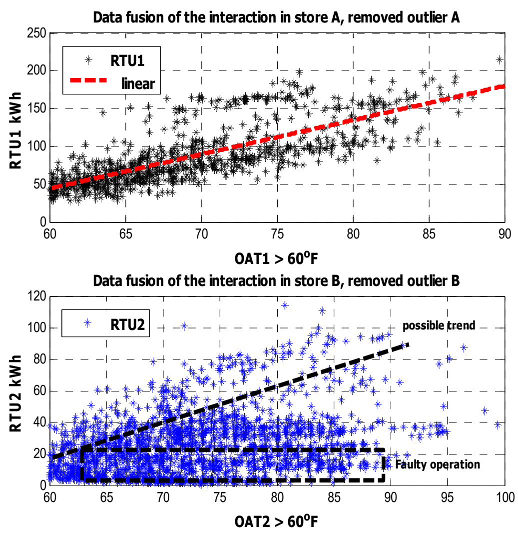

Fig. 6 Outlier removing of RTU kWh in Step 4

4.2 Z-score Implementation

Before computing Z score, an average and SD value of each data set are computed then are used to convert the data in Z score. Keeping z between -3 and 3, a new data set is determined and stored for analyzing energy interactions. Till now, step 1 to step 4 can be illustrated in Fig. 6.

In store A, it is clear that the linear line can be fitted from this data set relating to RTU operations versus OAT. Comparing store A performance to store B, there are still faulty operations because several data points are lower than 20 kWh. There are at least two possible causes: 1) faulty sensors lead to the trend deviation and 2) RTUs are over designed severely; they lead to uncoordinated control of compressor operations and frequent compressor cycling at part load conditions. Therefore, refrigeration power consumptions are helpful to further decrease these fault effect based on IARH considerations.

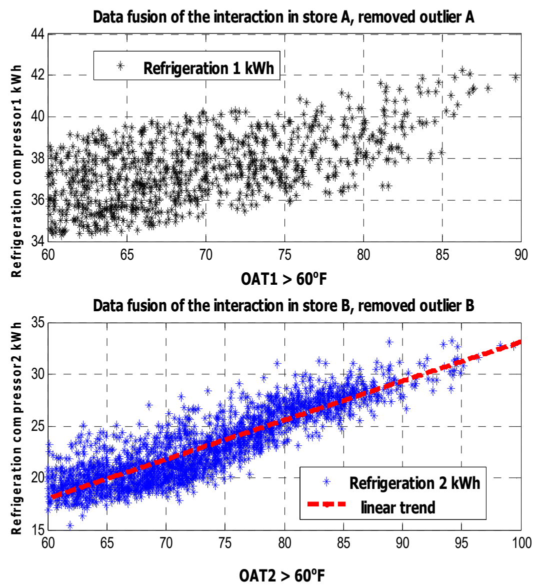

Fig. 7 Outlier removing of refrigeration kWh in Step 4

4.3 Data Improvement via Consistent IARH Range

In Fig. 7, after removing outlier in store A and B, the relations of OAT and refrigeration power consumptions are almost linear function as one of the driving force conditions; however, with different IARH conditions, the data are still coupled with the IARH effect. Therefore, in step 5, the consistent IARH ranges - 30 – 40%, 41 – 50% and 51 – 60% are selected to significantly decouple IARH impact from OAT based on humidity control function s. These fixed IARH values can be utilized to identify some malfunctions as shown in Fig. 8. The two IARH ranges (30 - 40% and 41 – 50% IARH) of RTU2 are compared to an unconstrained IARH range of the RTU2. It can be seen that the different range operations significantly result in different data qualities.

With the fixed IARH ranges and abnormal operations of the RTU 2 in Fig. 6, these two constraints can be used for improving the data fusion of RTU 2 as demonstrated in Fig. 9. Between 50 and 60% IARH, there are two-stage operations of RTUs leading to artificial two linear lines, whereas the baseline power consumptions of the multiple fan operations are around 20 kWh. However, RTU operations still maintain suitable zone conditions in store B resulting in well-conditioned relation between refrigeration compressor kWh and IARH2 at the same fixed range. This multi-step data fusion is very helpful to potentially analyze abnormal operations of HVAC&R equipment.

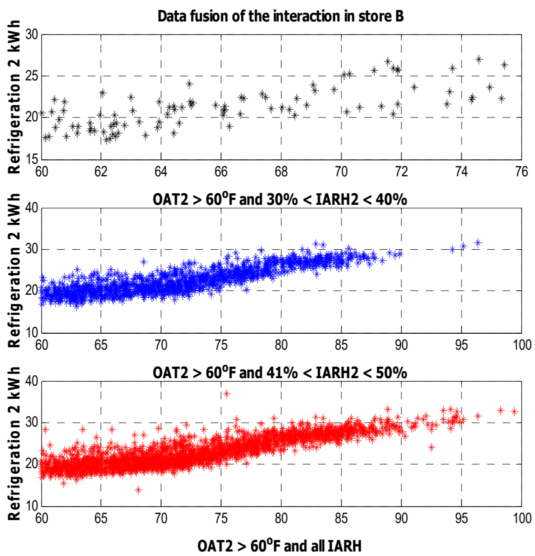

Fig. 8 Data Isolation via fixed IARH range in store B

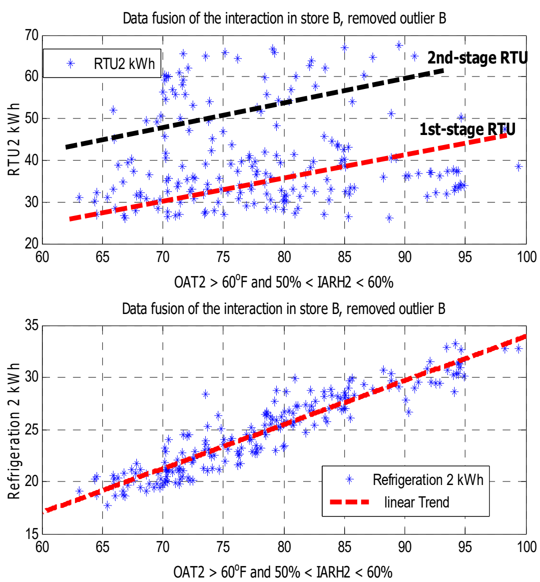

Fig. 9 Data Improvement in store B

5. Conclusion

The study proposes the data fusion for improving supermarket data quality obtained from BACnet protocol. The methodology is developed for investigating equipment operations through power consumption and measured variables without using fault-free data since laboratory data of a supermarket are cost-prohibitive. Meanwhile, in a field test, common faults are gradually developing in routine operations due to improper commissioning, installations and preventive maintenances. To overcome the barrier, multi-step data fusion technique is proposed to isolate outlier by using z-score, equipment control function and fixed IARH range to decouple IARH effect from refrigeration power consumptions. This technique can be further used to develop fault-free data for constructing simplified fault-free models in FDD applications and to improve the existing energy interaction performance.

6. Acknowledgement

The authors are deeply grateful to the Sripatum University and Thailand Research Fund (TRF) – MRG5980208 for financial support to further develop innovative technologies based on this work in near future.

7. References

[1] DOE (2010), Energy efficiency and renewable energy, energy efficiency trends in resi-dential and commercial buildings, Buildings Energy Data Book, August 2010.

[2] Woradechjumroen, D., Li, H. and Yu, Y. (2014). Energy interaction among HVAC and supermarket environment. *International Journal of Civil, Architectural, Structural and Construction Engineering* 2014; Vol 8 (12): 1119-1126.

[3] Kosar, D. and Dumitrescu, O. (2005). Humidity effects on supermarket refrigerated case energy performance: A Database Review. *ASHRAE Transactions*. Vol. 111 (1).

[4] Woradechjumroen, D., Li, H. and Yu, Y. (2016) Soft-repair technique for solving inherent oversizing effect on multiple rooftop units in commercial buildings. *Building and Environment* 2016; Vol. 108: 47- 62.

[5] Woradechjumroen, D. and Li, H. (2015). Smart building solutions: Introduction to methodologies. *Advances in Computer Science Research*. 2015; Vol 4: 980-982.

[6] Yu, D., Li, H., Ni., L and Yu, Y. (2011). An improved virtual calibration of a supply air temperature sensor in rooftop air conditioning units, *HVAC&R Research 2011*: Vol. 17 (5): 798-812.

[7] Stum, K. (2006). Sensor accuracy and calibration theory and practical application: National Conference on Building Commissioning, April 19-21.

[8] Henderson, H. (1998). The Impact of part load air conditioner operation on dehumidification performance: validating a latent capacity degradation model, presented at IAQ & Energy. New Orleans, LA, Oct. 1998.

[9] Woradechjumroen, D., Y. Yu, H. Li, D. Yu, and H. Yang. (2014). Analysis of HVAC system oversizing in commercial buildings through field measurements. *Energy and Buildings*, 69, P131–143.