

Article

A Temperature-Risk and Energy-Saving Evaluation Model for Supporting Energy-Saving Measures for Data Center Server Rooms

Kosuke Sasakura *, Takeshi Aoki, Masayoshi Komatsu and Takeshi Watanabe

NTT FACILITIES INC, Kotoku, Tokyo 135-0007, Japan; aokita28@ntt-f.co.jp (T.A.); komats36@ntt-f.co.jp (M.K.); watana25@ntt-f.co.jp (T.W.)

* Correspondence: sasaku24@ntt-f.co.jp; Tel.: +81-3-5669-0753

Received: 1 September 2020; Accepted: 4 October 2020; Published: 7 October 2020



Abstract: As data centers have become increasingly important in recent years their operational management must attain higher efficiency and reliability. Moreover, the power consumption of a data center is extremely large, and it is anticipated that it will continue to increase, so energy saving has become an urgent issue concerning data centers. In the meantime, the environment of the server rooms in data centers has become complicated owing to the introduction of virtualization technology, the installation of high-heat density information and communication technology (ICT) equipment and racks, and the diversification of cooling methods. It is very difficult to manage a server room in the case of such a complicated environment. When energy-saving measures are implemented in a server room with such a complicated environment, it is important to evaluate “temperature risks” in advance and calculate the energy-saving effect after the measures are taken. Under those circumstances, in this study, two prediction models are proposed: a model that predicts the rack intake temperature (so that the temperature risk can be evaluated in support of energy-saving measures implemented in the server room) and a model that evaluates the energy-saving effect (in relation to a baseline). Specifically, the models were constructed by using machine learning. The first constructed model evaluates the temperature risk in a verification room in advance, and it was confirmed that the model can evaluate the risk beforehand with high accuracy. The second constructed model—“baseline model” hereafter—supports energy-saving measures, and it was confirmed that the model can calculate the baseline (energy consumption) with high accuracy as well. Moreover, the effect of proposal process of energy-saving measures in the verification room was verified by using the two proposed models. In particular, the effectiveness of the model for evaluating temperature risk in advance and that of a technology for visualizing the energy-saving effect were confirmed.

Keywords: temperature prediction; energy simulation; baseline; machine learning; data center; server room; temperature environment; energy saving; continuous and reliable operation

1. Introduction

Information and communications technology (ICT) systems have become important tools for building the infrastructure that supports social life. Against that background, the role of data centers—which manage that information—is becoming more important [1]. Moreover, the power consumption of a data center is huge, and it is predicted to increase from now onwards; accordingly, reducing that power consumption has become an urgent issue [1,2]. In the meantime, owing to factors such as cloud computing, virtualization technology, high heat-generation density of ICT equipment, and diversification of cooling methods, the environment surrounding a data center is becoming ever more complicated [3,4]. Even with such a complex environment, a data center must be operated at

higher efficiency and higher reliable. Under such circumstances, various energy-saving measures for data centers are being studied. Among them, energy saving related to cooling, which accounts for about 40% of power consumption of a data center, is drawing attention [5]. However, implementing energy-saving measures in an operating server room faces two major obstacles.

One obstacle is proper temperature management. In particular, if a “service-level agreement” (SLA) between the operator of a data center and the users of that data center has been signed, the rack intake air temperature must be kept below a certain temperature. As a result, the demand for proper temperature control is a very high [5]. However, it is very difficult to properly control temperature in a complicated environment in which things dynamically change both temporally and spatially. In such an environment, if energy-saving measures related to cooling are implemented without evaluating the temperature risk in advance, so-called “hot spots” may occur and cause deterioration of service quality or suspension of services. In addition, it can be seen that temperature management is important because there is a demand for sudden rise suppression measures such as aisle containment and thermal storage systems for these problems. Tsukimoto et al. evaluate the temperature-rise suppression when CRAC is outage [6]. Lin et al. constructed a model to calculate the temperature rise after the CRAC was stopped and showed the possibility that the calculated result could be used for cooling and power design [7]. Garday et al. evaluate the effect of thermal storage when DC is outage [8]. Tsuda et al. and Niemann et al. conduct a comparative experiment on the aisle container method from the viewpoint of the influence of the intake temperature of the ICT equipment [9,10]. Therefore, it is very important to consider thermal environment, and we focused on the technology that predicts the rack intake temperature with high accuracy in advance, which realizes appropriate temperature management in the server room.

Several approaches to predicting temperature can be taken. One approach is to calculate the temperature at representative points by using the heat-balance equation. Although this approach can predict the temperature of representative points, it is difficult to predict the temperature of each rack unit in a server room. The other approach is to use simulation tool (for example, a computational fluid dynamics (CFD) tool for thermal-fluid simulation). In previous studies using CFD include the following. Sakaino uses CFD to set the CRAC based on the load balance of the ICT equipment and to predict the temperature environment in the server room [11]. Winbron et al. and Lin et al. use CFD to understand the effect of aisle containment on the temperature environment of the server room [12,13]. Jinkyun et al. use CFD to predict the change of temperature environment in the server room when a power outage or CRAC failure is occurred. They propose a method to use the result in the design of emergency power supply equipment [14]. Zavřel et al. use the simulation tool TRANSYS [15] to calculate the temperature rise when CRAC is stopped, and report on the consideration of the backup system of the power supply equipment and thermal storage to maintain the intake temperature of the ICT device properly [16]. Kummert et al. use the same simulation tool TRANSYS and propose a control method for a cooling system that can maintain an appropriate temperature environment when the CRAC is stopped [17]. Thus, This CFD and some simulation tool can predict the temperature of the entire server room. However, these approaches are difficult to adapt to multiple server rooms because modeling and calculation based on some simulation tools take a considerable amount time. What’s more, owing to the characteristics of a server room, the simulation results and measurement values often deviate. Tuning of simulation tool parameters for each room is therefore required. To solve these problems, a “rack intake temperature prediction model”—which is constructed by machine learning using data from a server room—that performs self-learning with high accuracy is proposed as described hereafter.

The other obstacle is to understand the effects of energy-saving measures. When implementing energy-saving measures, it is important to be able to quantitatively understand the extent of the energy-saving effect of each measure. If the energy-saving effect cannot be properly determined, the user will not know how effective the measure is; thus, the motivation for implementing that energy-saving measure will be reduced, and the measure itself may not be implemented at all. It is

therefore considered important to calculate the power consumption in the server room (“baseline” hereafter) when energy-saving measures are not implemented. As for energy-service-company (ESCO) projects targeting building equipment, the energy-saving effect may be evaluated by using statistical values (such as average and minimum) of past data (the previous year’s data, for example) as the baseline [18]. Jian et al. estimate yearly power consumption of cooling system by using maximum and minimum value of some facilities [19]. However, it is not desirable to use statistical values such as average, minimum or maximum values of past data as a baseline in an environment that changes fluidly in time and space (like a server room), since it is unknown whether the effect can be properly evaluated. In addition, when evaluating energy saving measures, it is preferable to estimate the baseline on a monthly, daily, or hourly basis rather than on an annual basis. Other approach is energy simulation (For example, CFD is used in DC, TRANSYS [15] and EnergyPlus are used in buildings). Maurizio et al. use TRANSYS and calculate the energy consumption in the electric steelmaking plants to evaluate energy and cost by selecting cooling methods [20]. Chao et al. use EnergyPlus to calculate energy consumption for three type building (large office, small office, residence) [21]. Cheng et al. report the results of calculation to use EnergyPlus based on BIM [22]. Stefan et al. combine EnergyPlus and CFD to calculate COP of cooling system in the building [23]. However, as with temperature prediction, it takes a lot of time to build a model and set boundary conditions, and its versatility is low. There is also a technique for formulating a baseline by regression equation using historical data for the issues related to energy simulation. Chao et al. constructed a regression model and confirmed that energy in large office, small office, and residence can be predicted with high accuracy [21]. In addition, Massimiliano et al. constructed a multiple regression model based on the outside air temperature, etc., and showed that the power consumption in the building can be predicted [24]. From these studies, it is considered that the baseline can be estimated quickly and with high accuracy by using a model constructing from the past data. Also, baseline technology is often applied to office buildings and residence, and there are not many studies on baselines for CRAC in server room. In the present study, we therefore propose a model that calculates the baseline for CRACs in a server room that is suitable for their current operation. The model is constructed by using machine learning with data from a server room, and it can perform self-learning with high accuracy.

When implementing energy-saving measures for data centers, it is important to execute the following four steps appropriately: (i) Planning of the measures, (ii) evaluate the temperature risk before implementing them, (iii) implement them, and (iv) evaluate their energy-saving effects (Figure 1). Accordingly, in this study, two models are proposed: a “rack intake temperature prediction model” that evaluates the temperature risk before implementing the energy-saving measures; and a “baseline model” that supports the evaluation of the energy-saving effect. The accuracy of each model was evaluated by experiments in a verification room. Energy-saving measures were implemented by applying these highly accurate models, and the effectiveness of each model was confirmed.

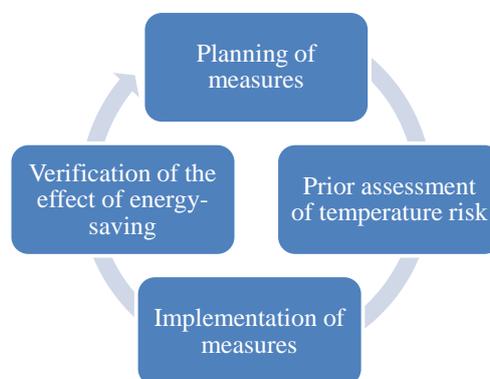


Figure 1. Four-step process of energy saving in a data center.

2. Verification Data Conditions

2.1. Verification Room

The configuration of the verification room is outlined in Figure 2, and its specification is listed in Table 1. The verification room is equipped with two computer room CRACs (CRACs), two “task-ambient” CRACs (racks A3 and B5), and 26 racks (four rows, A to D) for ICT equipment. The target of the energy-saving measures in this study is two of the CRAC (excluding the task-ambient type ones). The verification room is configured in the “cold-aisle containment” manner [8].

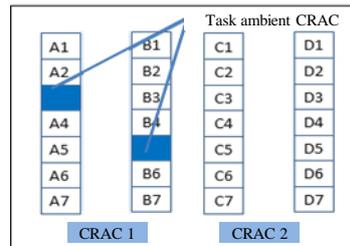


Figure 2. Floor plan of verification room.

Table 1. Specifications of verification server room and CRAC.

Item	Data
Room size (m ²)	140
Number of racks for ICT equipment	26
Number of CRACs	2
Number of task-ambient CRACs	2
Cooling capacity of CRAC (kW)	45

2.2. Verification Data

In recent years, data center infrastructure management (DCIM) systems, which support efficient operation by integrated management of various facilities and equipment in data centers, have become a focus [25,26]. Also, in previous study, we have summarized concepts and examples of effects related to DCIM [27]. We use the DCIM as an onsite data collecting system installed inside a server room (Figure 3). The data collecting system is connected to the CRAC, rack intake temperature sensors, and power-distribution units (PDUs) for the racks via local area network (LAN). This system collects and stores data from the connected devices and devices. The data-collection interval for the CRACs and temperature sensors is one minute, and that for the rack PDUs is five minutes. The rack intake temperature prediction model uses the data accumulated by the data-collection system. In this study, values aggregated over a 30-min average were used. Rack intake temperature is defined as the value taken by the temperature sensor installed on the rack surface at a height of 1.5 m (Figure 4).

The operation data listed in Table 2 was used for constructing and evaluating the rack intake temperature prediction model and the baseline model for the CRACs. The models were evaluated during the period from 21 October to 23 December, 2019; that is, the effectiveness of each model was verified by simulating an energy-saving measure that changes the setting of the CRAC return temperature.

Table 2. Data-set periods used with each model.

No	Verification Model	Period
1	Rack intake temperature	1 April to 30 April, 2016
2	prediction model	1 May to 31 October, 2017
3	Baseline model	1 May to 31 December, 2017
4	Both models	21 October to 23 December, 2019

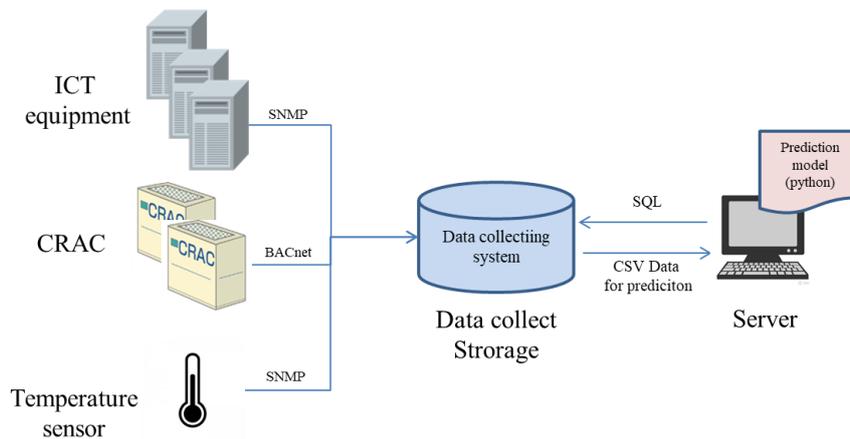


Figure 3. System configuration.

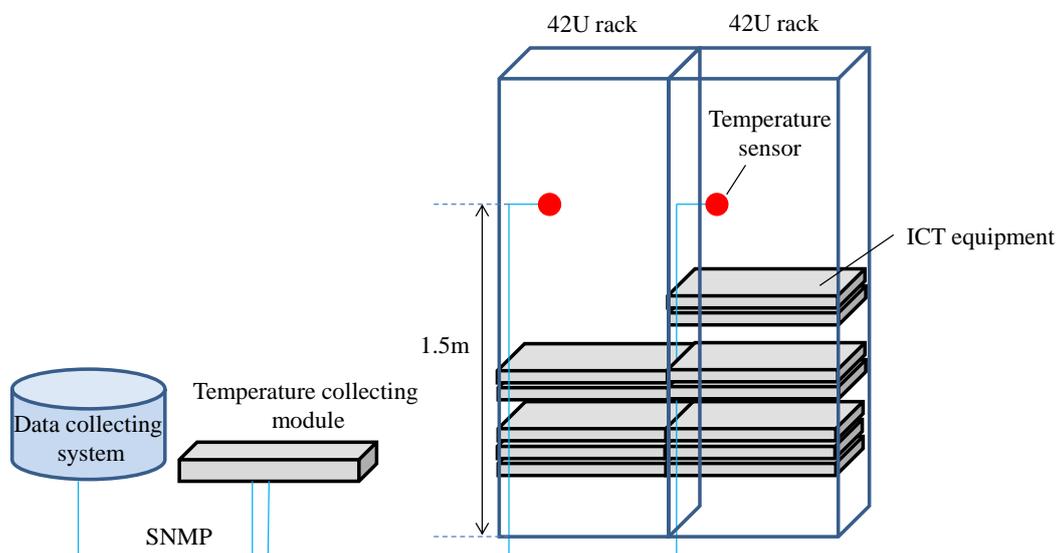


Figure 4. Installation of temperature sensor and way to get data.

3. Construction of Prediction Model

The rack intake temperature prediction model and the model for calculating the energy-consumption baseline for the CRACs are described hereafter.

3.1. Aim of Each Construction Model

In our research, as described in Section 1, we propose energy-saving measures in data center with a load that is highly fluid in time and space, including the prior assessment temperature risk and the verification of the effect of energy-saving measures, in response to the issue of energy-saving in data centers. As a technical issue to realize the proposal, we think that the technology to predict the rack intake temperature and calculate the baseline of CARC is necessary, and this chapter describes the construction method for the two models.

Regarding prediction rack intake temperature, there is a problem that in the previous research it took a lot of time to build a high accuracy model using CFD. Therefore, we construct and verify models which are multiple machine learning methods that can be constructed by self-learning using past data, and our goal is to construct highly accurate models. In addition, regarding the baseline, there is a problem that it is difficult to apply the baseline of CRAC in data center with a load that is highly fluid in time and space by using the past statistical values (max value, minimum value) that are often tackled in the past studies. We aim to build and verify a highly accurate model of machine

learning that can be built by self-learning using past data, using multiple methods selected from the viewpoint of AI's explanatory power.

3.2. Outline of Prediction Model of Rack Intake Temperature

The model that predicts rack intake temperature is described first.

3.2.1. Construction of Model for Predicting Rack Intake Temperature

As for this prediction model, rack intake temperature in one hour (or in 30 min) is the target variable. The explanatory variables used in the prediction are listed in Table 3. The prediction model divides the data set into a learning period and a verification period, and the model constructed by using the data set for the learning period is evaluated with the data set for the verification period. In this verification, data acquired up to n days before the verification period was set as the learning-period data.

Table 3. Explanatory variables used in the model for predicting rack intake temperature.

No.	Explanatory Variable
1	CRAC power consumption
2	CRAC COP
3	CRAC cooling capacity
4	Return temperature of CRAC
5	Supply temperature of CRAC
6	Power consumption of entire server room
7	Power consumption of each rack

3.2.2. Methods Used by the Model for Predicting Rack Intake Temperature

In a previous study, we were studying a model that predicts rack intake temperature after 30 min due to changes in ICT equipment in the server room and confirmed that it is possible to predict with high accuracy when using two methods (gradient boosting decision tree (GBDT) and a state space model) [28]. Based on the results, three methods, which are linear regression that are commonly used and two methods of the previous study, were selected as candidates. (Table 4).

Table 4. Methods selected as candidates for rack intake temperature prediction model.

No.	Method
1	Linear regression
2	Gradient-boosting decision tree (GBDT)
3	State-space model

First, the prediction model using each method was verified by using a dataset for April 2016, and the accuracy of each method was evaluated, and the features of each method were identified. According to the results of the evaluations, the methods were narrowed down, further verified and evaluated in detail by using a dataset acquired from May to October 2017, and the model was selected. When the gradient-boosting decision tree (GBDT) was selected as the machine-learning method, the library "XGBoost," which is a Python implementation of scikit-learn, was used. The grid-search method was used with parameters set as follows: learning rate: 0.1; lower limit of loss reduction due to addition of leaves: 0; maximum depth of tree: 3; ratio of randomly sampled samples (data): 1; and ratio of columns randomly extracted from each decision tree: 1. As for other parameters, the default values in the library were used. In the periods other than No. 1, the parameters optimized by the grid search were also optimized by the same method. When the state-space model was selected as the machine-learning method, the python library "statsmodels" was used. Parameters were estimated by using a Monte Carlo filter and pseudo-Newton method.

3.2.3. Method for Evaluating the Model for Predicting Rack Intake Temperature

The rack intake temperature prediction model was evaluated by using the following four evaluation indices:

1. *Correlation coefficient (R)*: R expresses the explanatory power of the predicted value of the objective variable.
2. *Correct answer rate*: The ratio of the number of predicted values within ± 0.5 °C of the measured value to the total number of predicted values
3. *Root-mean-square error (RMSE)*: The accuracy of the three machine-learning methods is evaluated in terms of RMSE, which is a commonly used index for numerical prediction.
4. *Maximum peak error*: As for predicting server-room temperature, maximum peak error is significant if the actual measured value and the predicted value deviate greatly. The error by which the actually measured value is larger than the predicted value is therefore defined as maximum peak error.

3.3. Outline of Baseline Model of CRAC

The baseline model for CRAC is overviewed as follows.

3.3.1. Construction of a Baseline Model for CRAC

The purpose of the baseline model is to determine the effect of energy-saving measures. The objective variable of the model is therefore the power consumption of CRAC when the energy-saving measure is not implemented. The baseline model calculates power consumption in real time when the energy-saving measure is not implemented; that is, it is not the model that predicts the future from that time point. The explanatory variables used are listed in Table 5. The dataset is divided into a learning period and a verification period, and the model constructed by using the learning-period dataset is evaluated by using the verification-period dataset. In this verification, the data collected up to n days before the verification period was set as the data for the learning period.

Table 5. Explanatory variables used by the baseline model.

No.	Explanatory Variable
1	CRAC cooling capacity
2	Outside-air temperature
3	Power consumption of each rack

3.3.2. Methods Used by Baseline Model

In recent years, “explainable” AI (XAI), which can explain prediction results by machine learning and processes in humans, has received increasing attention [29,30]. This explainable model allows the user to explain and understand the result calculated by the model. It is essential that the model for determining the effect of energy-saving measures targeted in this study should be explainable to scholars with specialized skills and operators who carry out practical work. In addition, as in the ESCO business, compensation may be paid for energy-saving effects; accordingly, from the standpoint of data-center businesses too, it is desirable to be able to confirm the likelihood of effects. The baseline model should therefore have high explanatory power. In this study, therefore, a total of three methods were selected: (1) linear regression (which is considered to be highly explanatory); (2) a decision-tree and, (3) gradient-boosting decision tree (which is an application of a decision-tree model that is expected to improve accuracy). Also, a state-space model. In addition, the state-space model (time-series model) used as a candidate for the rack intake temperature prediction model in this model has the feature that it is updated sequentially with data acquired at the preceding n time points. Therefore, when calculating the baseline without energy-saving measures implemented, it is a problem that prediction accuracy

decreases because the prediction result is used for the next prediction; thus, subject was excluded from the method used by the baseline model.

3.3.3. Method for Evaluating Baseline Model

Accuracy of the baseline model was evaluated by using evaluation indices 1, 3, 4 and 5. This is because it was considered inappropriate to use evaluation index 2 (correct-answer rate) because the power consumption of the CRAC has a change significantly compared to temperature. Also, in the baseline model, NMBE is added as indices. This is a commonly used index when performing energy simulations [31,32]:

1. *Correlation coefficient (R)*: R expresses the explanatory power of the predicted value of the objective variable.
3. *Root-mean-square error (RMSE)*: The accuracy of the three machine-learning methods is evaluated in terms of RMSE, which is a commonly used index for numerical prediction.
4. *Maximum peak error*: As for predicting server-room temperature, maximum peak error is significant if the actual measured value and the predicted value deviate greatly. The error by which the actually measured value is larger than the predicted value is therefore defined as maximum peak error.
5. *Normalized Mean Bias Error (NMBE)*: NMBE is a normalization of the MBE index that is used to scale the results of MBE, making them comparable. This index is used by IPMVP.

4. Results

4.1. Evaluation of Accuracy of Temperature-Prediction Model

4.1.1. Primary Evaluation of Prediction Model and Narrowing Down of Prediction Methods

A model for predicting the rack intake temperature was constructed using each method under the assumption that the verification period was the last week of April 2016 and the learning period was the previous 21 days, and the accuracy of the model was evaluated. Evaluation values when each method was used are listed in Table 6. According to these values, evaluation indexes 1 to 3 used the average of evaluation values for all racks, and evaluation index 4 used the maximum value of the peak error for all racks. According to these results, it is clear that GBDT and the state-space model give similar values for evaluation indexes 1 to 4, meaning that the predictions by those methods achieve high accuracy. On the contrary, although linear regression gives similar correct-answer rate and peak error to those given by GBDT and the state-space model, correlation coefficient R and RMSE are low, so it can be concluded that the accuracy of linear regression is low compared to the other two methods. In light of that result, hereafter, GBDT and the state-space mode are focused on and investigated in detail.

Table 6. Evaluation index of each model (April 2016).

Method	Evaluation Index			
	R	Correct-Answer Rate	RMSE	Max Peak Error
Linear regression	0.34	0.86	0.36	0.97
GBDT	0.81	0.99	0.11	1.07
State-space model	0.94	0.99	0.10	1.04

4.1.2. Secondary Evaluation of Prediction Model and Determination of Prediction Method

The two methods narrowed down in Section 4.1.1 are summarized in Figures 5 and 6, and Table 7. As for the secondary evaluation conducted from May to October 2017, the learning and verification periods were taken as one month each, and the data of the month previous to the verification-target month was set as the learning period (Table 8). Evaluation values that compare the measured rack

intake temperature with the temperature predicted by using each prediction model are listed in Table 9. Evaluation indexes 1 to 3 are average evaluation values for all racks, and evaluation index 4 is the maximum evaluation value for all racks, and these values were obtained by tabulating the prediction results for each month. It is clear from these results that the prediction values are extremely high and the prediction achieves high accuracy regardless of which of the two prediction models is used. It is also clear that the evaluation values (R and correct-answer ratio) given by the state-space model are higher than those given by GBDT, although they are only slightly higher over the entire period.

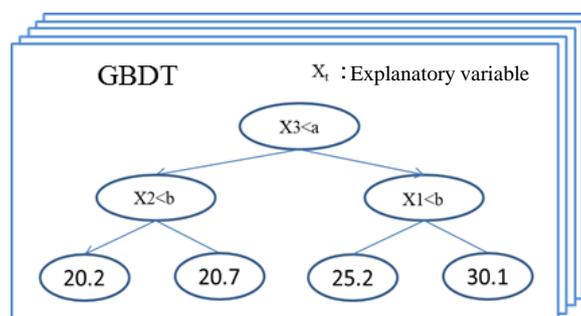


Figure 5. Overview of GBDT.

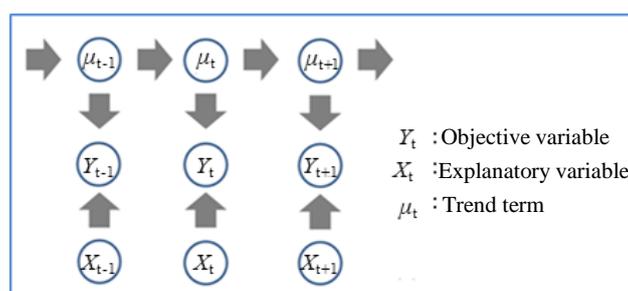


Figure 6. Overview of state-space model.

Table 7. Methods selected as a candidate for rack intake temperature prediction model.

Method	Explanation
GBDT	As an ensemble learning method using decision trees, a prediction method used in regression and classification problems [33]
State-space model	Prediction method used for time-series problems [34]

Table 8. Relationship between learning period and evaluation period.

2017					
May	June	July	August	September	October
Learning	Evaluation Learning	Evaluation Learning	Evaluation Learning	Evaluation Learning	Evaluation

Table 9. Evaluation indices for each method (May to October 2017).

Method	Evaluation Index			
	R	Correct-Answer Rate	RMSE	Max Peak Error
GBDT	0.82	0.62	0.35	2.77
State-space model	0.98	0.99	0.14	2.20

Next, the accuracy of the two methods, GBDT and state-space model, was further examined on the basis of temporal changes in the measured and predicted temperatures. Measured rack intake temperature for rack A1 in July 2017 and time-series changes of values predicted by each prediction model are respectively shown in Figure 7. According to these figures, the predicted temperature values follow the time-series changes in measured values closely. Measured rack intake temperatures for rack A7 in July 2017 and time-series changes of temperatures predicted by each prediction model are respectively shown in Figure 8. These figures show that the error between the measured temperatures and the temperature predicted by the model using GBDT increased on prediction date 5 July 2017. It is considered that this increased error is due to two reasons: (i) the effect of the intake temperature of rack A7 at this time point during the learning and verification periods being the lowest of all the racks and (ii) in the case of GBDT, if the input fluctuates significantly, the prediction result cannot be handled and becomes worse. On the contrary, as shown in Figure 8b and Table 9, in the case of the state-space model, which gives higher evaluation values than those of GBDT, the result obtained at the previous time point is reflected in the prediction model, and the model is updated sequentially; as a result, the ability to follow extrapolation is high, and is considered that good prediction results were obtained.

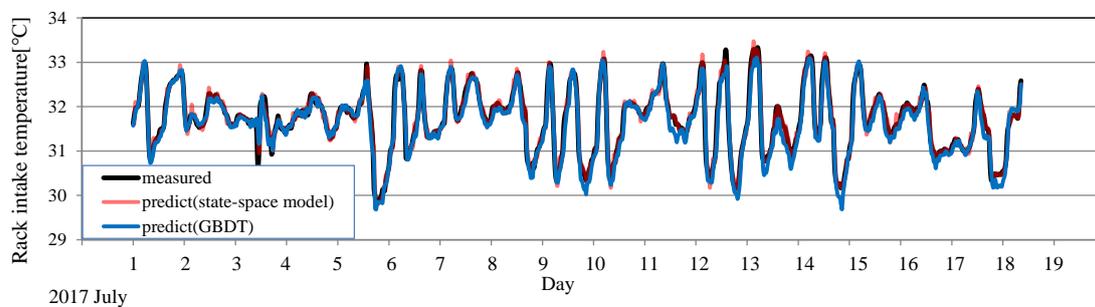


Figure 7. Comparison of measured rack intake temperature in rack A1 and temperatures predicted by prediction model using GBDT and state-space model.

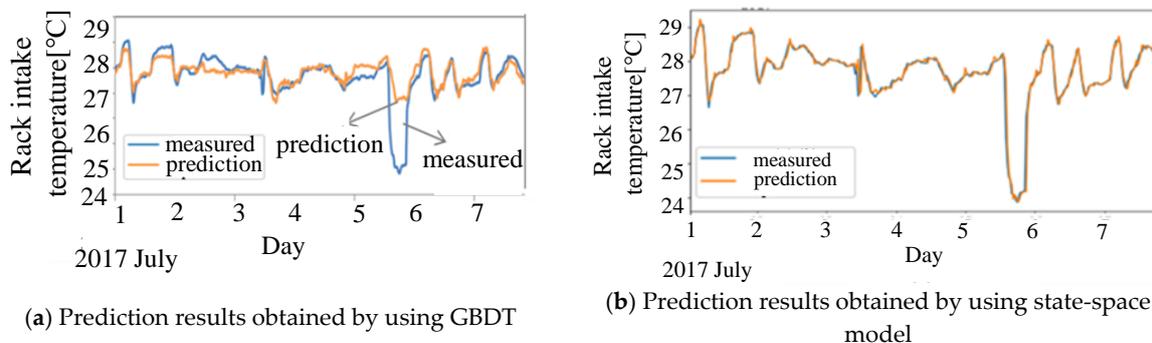


Figure 8. Time-series change of rack intake temperature in rack A7.

4.1.3. Detailed Evaluation of the Determined Prediction Model

To understand the features of the prediction model, the effects of each explanatory variable and learning period on prediction accuracy are considered hereafter.

1. Effect of explanatory variables on accuracy

When building a predictive model, the choice of explanatory variables is paramount. This is because not only the choice of explanatory variables affects accuracy but also the time required to build a prediction model changes significantly. Therefore, to study the effect of each explanatory variable on the accuracy of the prediction model, change in RMSE was observed when the explanatory variables were changed in two major ways. RMSE was used because it makes it possible to quantify and evaluate the effect of the explanatory variables on the accuracy of the model by evaluating how

much a predicted value has changed from the actual value during a certain period. One model was verified when one explanatory variable was removed from all the explanatory variables, and the other model was verified when only one explanatory variable was used as an explanatory variable. In period No. 2 (May 1 to October 31, 2017; see Table 2), the effect of the explanatory variables was evaluated in terms of the value of RMSE, and the results are listed in Table 10. It can be seen that when one explanatory variable is removed from all explanatory variables, the RMSE value drops significantly only when the CRAC return temperature is removed from the explanatory variables. When the other explanatory variables were removed, RMSE does not change much. Next, as for the verification using only one variable as the explanatory variable, the best RMSE was obtained when only CRAC return-air temperature was used as the explanatory variable. These results demonstrate that return temperature of the CRAC significantly influences prediction accuracy. Furthermore, the time required for building the model did not significantly change when the explanatory variables were changed. This is probably because (i) the explanatory variables used in this study were not large in number and (ii) increasing or decreasing the number of variables had a small effect on building the model.

Table 10. RMSE when explanatory variable is changed.

Explanatory Variable Used	RMSE
Without CRAC power consumption	0.13
Without CRAC COP	0.13
Without CRAC cooling capacity	0.13
Without return temperature of CRAC	0.18
Without supply temperature of CRAC	0.13
Without power consumption of entire server room	0.13
Without power consumption of each rack	0.14
Only CRAC power consumption	0.22
Only COP of CRAC	0.21
Only CRAC cooling capacity	0.22
Only return temperature of CRAC	0.15
Only supply temperature of CRAC	0.21
Only power consumption of entire server room	0.22
Only power consumption of each rack	0.18
(Reference) when all variables are used	0.14

2. Effect of learning period on accuracy

When constructing a predictive model, it is paramount to consider the length of the learning period. This is because the learning-period length significantly influences not only prediction accuracy (as in the case of explanatory variables) but also the time associated with the construction of the prediction model. Therefore, in this section, to consider the influence of the learning period for the prediction model on the accuracy, the model was verified by changing the learning period while using the data for August 2017. The effect on prediction accuracy when the learning period was varied was evaluated in terms of the value of RMSE, and the evaluation results are listed in Table 11. As shown by the results in Table 12, RMSE does not significantly change when the learning period is varied. On the contrary, although the time required for learning increased by a maximum of about two times, that increased time did not cause any operational problems when only the server room in this study was targeted. These results indicate that when constructing a prediction model, it is possible to highly accurately predict temperature if data covering about one week is available. Note that the shortest learning period that can achieve high prediction accuracy was left for future study.

Table 11. Relationship between learning period and evaluation period.

2017							
May	June	July	August	September	October	November	December
learning	evaluation learning	evaluation learning	evaluation learning	evaluation learning	evaluation learning	evaluation learning	evaluation
							evaluation

Table 12. RMSE when the learning period is varied.

Learning Period	RMSE
7 days before the evaluation period	0.12
14 days before the evaluation period	0.13
21 days before the evaluation period	0.12
31 days before the evaluation period	0.14
61 days before the evaluation period	0.12
91 days before the evaluation period	0.13

4.1.4. Summary of Evaluation of Accuracy of Temperature Prediction Model

As explained hereafter, it was confirmed that the model for predicting rack intake temperature using the state-space model, which was narrowed down from the models constructed by multiple methods, can predict the temperature with high accuracy. Moreover, the effect of each explanatory variable used in the proposed model was evaluated, and the results of that evaluation demonstrate that the return temperature of the CRAC is important in regard to predicting the rack intake temperature using this state-space-based prediction model. After that, the effect of the learning period on the accuracy of the proposed model was evaluated, and the results of the evaluation performed in the verification room show that it is possible to predict temperature with high accuracy using about one week's worth of learning data.

4.2. Evaluation of Accuracy of Baseline Model

4.2.1. Primary Evaluation of Baseline Model and Narrowing Down of Prediction Methods

During the period from May 1 to December 31, 2017, the learning and verification periods were both set to one month, prediction models was constructed by using each method (linear regression, decision tree, and GBDT), and the prediction accuracy of each model was evaluated (Table 13). As for these values, evaluation indexes 1 and 2 are average evaluation values for each month, and evaluation index 4 is the maximum value for each month. According to these results, GBDT achieved the best values and linear regression achieved the worst values of evaluation indices (1), (3) and (4). Also, indices (1), (3) and (4) values given by the prediction model using the decision-tree method are worse than those given by GBDT. This finding is considered to be explained by the fact that in contrast to the decision-tree method, GBDT features added “boosting” for weighting, and the presence or absence of that boosting affects prediction accuracy. Regarding indices (5), linear regression is better than other methods, and we discuss about this indices and results in detail next section.

Table 13. Evaluation index for each model (April 2016).

Method	Evaluation Index			
	(1) R	(3) RMSE	(4) max peak error	(5) NMBE
Linear regression	0.82	1.00	4.75	0.61
Decision tree	0.85	0.61	3.11	1.00
gbdt	0.88	0.56	2.56	1.37

4.2.2. Secondary Evaluation of Baseline Model and Determination of Prediction Method

The evaluation values for each month during the period from May 1 to December 31, 2017 are listed in Tables 14–17 and plotted in Figures 9–12. Regarding the correlation coefficients listed in Table 14 and plotted in Figure 9, all methods show numerical values with high correlation, and it can be considered that prediction can be performed with high accuracy. Regarding the correlation coefficients listed in Table 14 and Figure 9, all three methods give numerical values with high correlation, so it can be considered that the prediction has high enough accuracy. Even so, for any month, GBDT did not give the worst value of the three methods; in fact, the results show that it often gave the best value. As for RMSE listed in Table 15 and plotted Figure 10, compared with the other methods, linear regression does not give the best value for any month, and its fluctuation per month is the largest. In addition, RMSE values given by the decision-tree method and GBDT are close. Although RMSE varies slightly in a similar manner to the correlation coefficient, the values given by GBDT are the best for several of the months. Regarding peak error listed in Table 16 and plotted in Figure 11, linear regression gives the best values for four of the months (July, October, November, and December). However, fluctuations of peak error per month are larger (minimum: 0.48 kW (July) and maximum: 4.75 kW (September)) than those of the other two methods. Although the peak errors of the decision-tree method and GBDT are close, GBDT gives the better value for many months. It is concluded from these results that linear regression is unsuitable as the baseline model because the evaluation indices related to prediction accuracy (i.e., RMSE and peak error) fluctuate from month to month. In addition, both the decision-tree method and GBDT give high correlation coefficients, and it is concluded that they can predict temperature with high accuracy. However, it is considered that GBDT, which on the whole gives better values than those of the decision tree, has superior robustness and is suitable as the baseline model. Accordingly, the model using GBDT was evaluated as explained hereafter.

Table 14. Correlation coefficient for each method per month. (bold and underline means a best value of each month).

Method	2017						
	June	July	August	September	October	November	December
Linear regression	0.77	<u>0.90</u>	0.87	0.87	<u>0.94</u>	<u>0.67</u>	<u>0.67</u>
Decision tree	0.84	0.85	0.84	0.84	0.94	<u>0.75</u>	0.91
GBDT	<u>0.90</u>	0.89	<u>0.89</u>	<u>0.88</u>	<u>0.95</u>	0.70	<u>0.94</u>

Table 15. RMSE for each method per month. (bold and underline means a best value of each month).

Method	2017						
	June	July	August	September	October	November	December
Linear regression	2.35	0.61	0.67	0.68	0.56	0.84	1.28
Decision tree	0.83	0.69	0.64	0.64	<u>0.44</u>	<u>0.37</u>	0.69
GBDT	<u>0.66</u>	<u>0.62</u>	<u>0.56</u>	<u>0.57</u>	0.54	0.41	<u>0.61</u>

Table 16. Peak error for each method per month. (bold and underline means a best value of each month).

Method	2017						
	June	July	August	September	October	November	December
Linear regression	3.85	0.48	4.71	4.75	0.68	0.51	0.40
Decision tree	2.20	1.10	<u>2.20</u>	2.37	2.78	<u>2.35</u>	3.11
GBDT	<u>1.64</u>	<u>0.75</u>	2.22	<u>2.20</u>	<u>1.14</u>	2.56	<u>2.16</u>

Table 17. NMBE for each method per month. (bold and underline means a best value of each month).

Method	2017						
	June	July	August	September	October	November	December
Linear regression	30.28	3.83	<u>-0.73</u>	<u>0.93</u>	6.86	10.27	15.25
Decision tree	-7.33	3.74	-2.22	-2.61	<u>-4.54</u>	<u>0.64</u>	<u>5.31</u>
GBDT	<u>6.78</u>	<u>3.69</u>	-2.14	-2.44	-7.73	-1.52	7.29

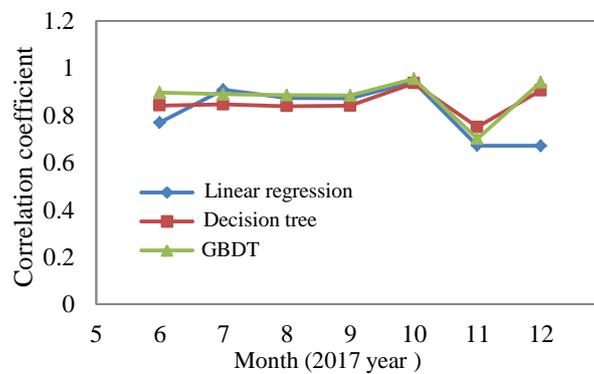


Figure 9. Transition of coefficient of determination for each method per month.

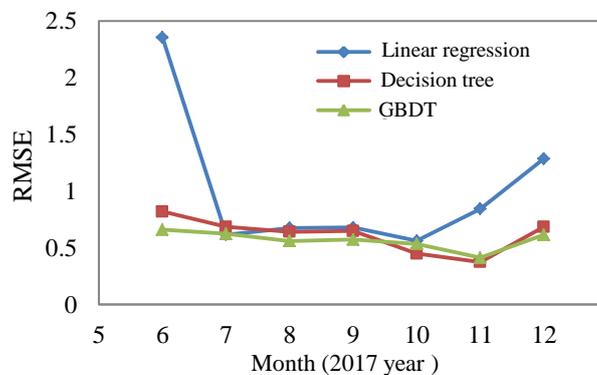


Figure 10. Transition of RMSE for each method per month.

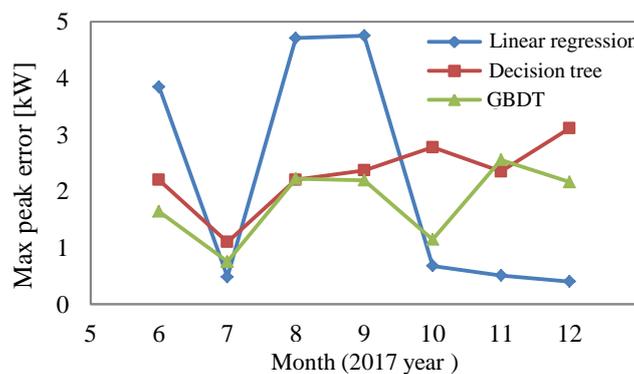


Figure 11. Transition of peak error for each method per month.

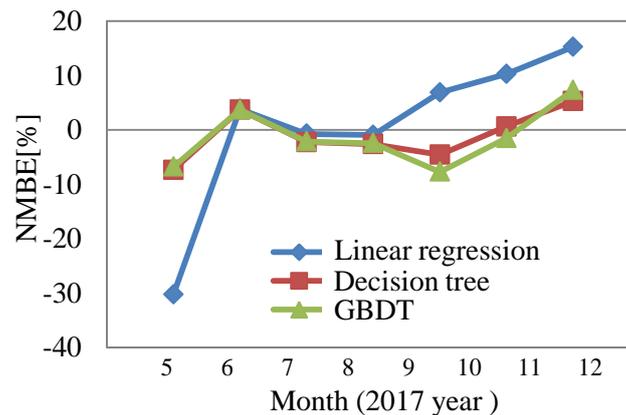


Figure 12. Transition of NMBE for each method per month.

Regarding the NMBE listed in Table 17 and Figure 12, since the recommended range per month proposed by IPMVP is $\pm 20\%$, it can be seen that the range is achieved by the decision-tree or GBDT. In addition, linear regression was the best value of NMBE when evaluated over the entire period. However, when analyzed monthly, the variation in linear regression was very large. Since NMBE does not use absolute values, the evaluation results vary greatly depending on the period. In the following, the RMSE, which uses absolute values, will be used to evaluate the error as a whole. For these reasons, the indices (1), (3), and (4) are used for evaluation thereafter.

4.2.3. Detailed Evaluation of Baseline Model

To understand the features of the baseline model, the effects of each explanatory variable and learning period on prediction accuracy are considered hereafter.

1. Effect of explanatory variables on prediction accuracy

The choice of explanatory variables is paramount when building a predictive model. That is because not only the learning period affects prediction accuracy but also the time required to build the model changes significantly with the explanatory variables chosen. Furthermore, when using the baseline model to verify the energy-saving effect of measures, it is important that estimated values that play no part in an energy-saving effect can be understood by or explained to the user. Accordingly, to investigate the influence of each explanatory variable on prediction accuracy of the baseline model, the “degree of importance” of the explanatory variables was estimated. Degree of importance means to what extent each explanatory variable contributes to improving prediction accuracy. Regarding the model when November 2017 was the learning period and December 2017 was the evaluation period, the results obtained by using the existing python library and calculated by using (`xgb.plot_importance(x)`) and default gain are shown in Figure 13. This result indicates that cooling capacity has the greatest effect on prediction accuracy. And although the degrees of importance of the other explanatory variables are less than half, they cannot be ignored and are considered to be important explanatory variables. Although it is a more complicated method than linear regression or decision-tree method, GBDT makes it possible to understand a part of the logic by showing the importance of the explanatory variable for the target variable as shown in the figure.

2. Effect of learning period on accuracy

With regard to the methods narrowed down as described in the previous section, the effect of length of learning period on prediction accuracy is considered hereafter. The evaluation period was fixed to one month (November or December 2017), and the learning period was changed from the last week to the last six months (Table 18). The evaluation indices for the accuracy during the one-month evaluation period (November or December 2017) when the learning period was changed are listed in

Tables 19 and 20. It can be concluded from the results in these tables that the relationship between the learning period and the evaluation indices is insignificant when the learning period is more than the previous three weeks. On the contrary, in the case of a very short learning period, such as one week, the results indicate that the accuracy values decrease.

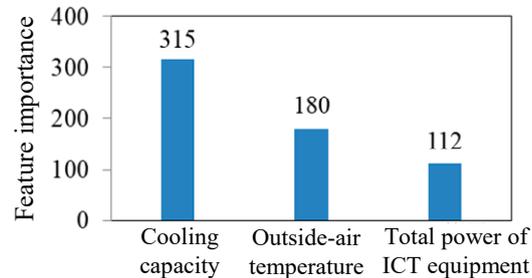


Figure 13. Importance of explanatory variables used in the baseline model.

Table 18. Explanatory variables used in the baseline model.

Learning Period No.	Explanatory Variable	Evaluation Period
1	Previous week	November and December, 2017
2	Previous two weeks	
3	Previous three weeks	
4	Previous month	
5	Previous two months	
6	Previous three months	
7	Previous four months	
8	Previous five months	
9	Previous six months	

Table 19. Evaluation indices when learning period is changed (evaluation month: November).

Evaluation Index	Learning Period No.								
	9	8	7	6	5	4	3	2	1
Correlation coefficient	0.67	0.53	0.60	0.59	0.52	0.70	0.74	0.52	0.51
RMSE	0.56	0.78	0.61	0.70	0.76	0.41	0.39	0.48	0.46
Peak difference	2.17	3.11	2.51	2.54	2.81	2.56	2.38	2.36	2.39

Table 20. Evaluation indices when learning period is changed (evaluation month: December).

Evaluation Index	Learning Period No.								
	9	8	7	6	5	4	3	2	1
Correlation coefficient	0.94	0.94	0.95	0.95	0.94	0.94	0.94	0.94	0.88
RMSE	0.59	0.62	0.60	0.58	0.62	0.62	0.60	0.63	0.99
Peak difference	1.36	1.31	1.61	1.64	1.96	2.16	2.11	1.90	1.64

4.2.4. Summary of Evaluation of Accuracy of the Baseline Model

It was confirmed that the baseline model for power consumption of an CRAC can calculate temperature with high accuracy by constructing a model using GBDT (which was narrowed down from three methods). Moreover, the effect of each explanatory variable used in the proposed baseline model was visualized by calculating the degree of importance of each explanatory variable. After that, the effect of the learning period on prediction accuracy of the proposed model was evaluated, and the results of that evaluation show that it is possible to predict temperature with high accuracy if there is about three weeks of learning data acquired in the verification room.

5. Verification of Effectiveness of the Proposed Model When Energy-Saving Measures Are Implemented

This section presents the results of a preliminary examination and effectiveness verification of energy-saving measures by using the rack intake temperature prediction model constructed and evaluated in the previous section and the baseline model for the CRACs.

5.1. Overview of Effectiveness Verification

As for the evaluation period from October to December 2019, the effect and effectiveness of the proposed model was verified by simulating an energy-saving measure that changes the setting of return-air temperature of the CRACs. The normal condition of the return temperature of the CRAC was set to 28 °C, and the return temperature setting was changed to 30 °C during the measure-implementation period (two weeks). Note that during the period excluding the verification experiment simulating the energy-saving measure, the accuracy of each model was verified, and it was confirmed that it the models can predict and calculate the temperature with high accuracy, as described in Section 4 (Tables 21 and 22). The results of an evaluation of temperature risk when the energy-saving measure was implemented and the evaluation results regarding the visualization of effect of the measure are presented in the following sections.

Table 21. Preliminary evaluation indices of rack intake temperature prediction model (8 November 2019 to 20 November 2019; sample number: 961).

Method	Evaluation Index			
	R	Correct-Answer Rate	RMSE	Max Peak Error
State-space model	0.99	0.98	0.14	2.37

Table 22. Baseline evaluation indices when return-temperature setting of CRAC is 28 °C (6 December 2019 to 10 December 2019; sample number: 192).

Method	Evaluation Index		
	R	RMSE	Max Peak Error
GBDT	0.91	0.26	0.49

5.2. Evaluation of Temperature Risk by Using Rack Intake Temperature Prediction Model

The change in rack intake temperature when the set value of the CRAC return temperature was changed to 30 °C was predicted in advance. Three explanatory variables were selected on the basis of three viewpoints: rack intake temperature at the previous time point, CRAC return temperature, and total power consumption of the server room. The rack intake temperature at the previous time point was selected on the basis of the characteristics of the state-space model using the value of the target variable at the previous time point. Return temperature was also selected because it is an important explanatory variable that affects prediction accuracy (as discussed in Section 4.1.4). Since the set temperature was set to 30 °C in this verification, return temperature was also set to 30 °C. About total power consumption, it is very difficult to predict the power consumption of each rack in DC. In previous studies, Mehdi et al. created a scenario to estimate the power consumption of multiple ICT devices, including the relationship between temperature and ICT equipment power consumption [35]. In this study, we decided to perform a simulation in the case of the maximum risk by adopting the maximum power consumption in the past week from the viewpoint of evaluating the risk. A model constructed as a learning period from November 8 to 29, 2019 was used, and the temperature for the period from November 21 to December 6, 2019, was predicted to determine, and the predicted temperature was compared with the measured value (Table 23). From this table, it can be seen that the calculation results show values very close to the average and median values. In addition, it is

clear that the error between the maximum value and the calculation result for all racks is within the range of 2.37 °C shown as the peak error in Table 23. This is because the max peak error value in the preliminary evaluation is considered to be the maximum error held by the model, so (calculation result + peak error) should be considered as the maximum rack intake temperature that may occur during the evaluation period.

Table 23. Relationship between the calculation results at each rack intake temperature and the maximum and average values of the measured values during the evaluation period.

Rack No	Calculation Results	Max Measured Value	Error (Max Measured Value—Calculation Results)	Average Measured Value	Error (Average Measured Value—Calculation Results)
A1	30.36	32.62	2.26	30.77	0.41
A2	29.52	31.45	1.93	29.52	−0.01
A3	30.63	32.61	1.98	31.09	0.46
A4	29.56	31.57	2.01	29.16	−0.40
A5	33.46	35.00	1.54	32.82	−0.64
A6	29.99	32.11	2.12	29.54	−0.45
A7	31.15	33.35	2.19	31.21	0.05
B1	30.86	32.89	2.03	31.09	0.23
B2	29.31	31.09	1.78	29.04	−0.28
B3	29.48	31.45	1.97	29.11	−0.37
B4	31.48	33.58	2.10	31.43	−0.04
B5	29.38	31.34	1.96	29.47	0.09
B6	31.76	34.06	2.30	31.61	−0.15
B7	29.13	31.46	2.33	29.03	−0.10
C1	30.38	32.67	2.29	30.86	0.47
C2	30.20	32.30	2.10	30.71	0.51
C3	32.85	34.92	2.08	33.15	0.31
C4	29.66	31.44	1.78	29.72	0.06
C5	29.05	30.76	1.71	29.14	0.09
C6	29.87	31.82	1.95	30.01	0.15
C7	28.80	30.53	1.73	28.87	0.07
D1	30.25	32.00	1.745	30.45	0.20
D2	29.99	31.77	1.78	30.20	0.21
D3	30.83	33.09	2.26	31.11	0.28
D4	30.01	31.49	1.48	29.84	−0.17
D5	29.53	31.14	1.61	29.34	−0.18
D6	28.88	30.73	1.85	29.03	0.15
D7	28.62	30.50	1.88	28.83	0.20

Since the maximum rack intake air temperature that may have been calculated during the preliminary study is larger than the maximum value during the evaluation period, it is considered useful for the preliminary study (Figure 14). Furthermore, for each rack, it is possible to predict which rack will become hot, so it is possible to understand the points to be noted when implementing energy conservation measures. It was confirmed the effectiveness of using this prediction model for the preliminary examination of temperature risk.

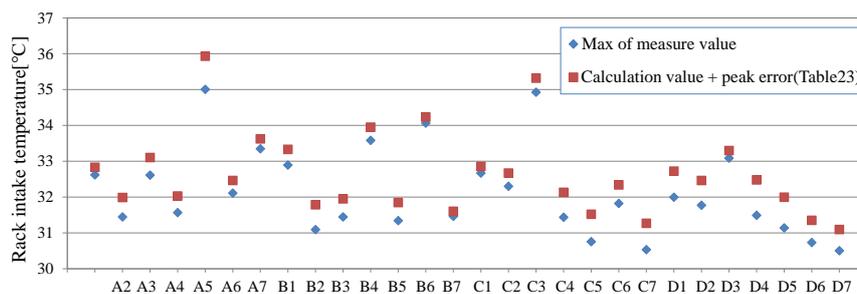


Figure 14. Relationship between calculation value + peak error and max of measure value for each rack.

5.3. Visualization of Energy-Saving Effect by Using Baseline Model

The energy-saving effect was verified when the return-air temperature of the CRAC, which was in the normal condition, was changed from 28 °C to 30 °C. The energy-saving effect calculated from the difference between measured power consumption and the value predicted by the baseline model when the return-air temperature of the CRAC was increased to 30 °C is shown in Table 24, and the time-series changes are plotted in Figure 15. It is clear that lowering the set temperature twice for about 14 days has a reduction effect of about 50 kW in terms of power consumption during that period. And it is clear from Figure 15 that in addition to the time-series changes with a similar tendency, the baseline shows a slightly higher proportion of high values.

Table 24. Energy-saving effect when return temperature of the CRAC is set to 30 °C (21 November 2019 to 6 December 2019; sample number: 706).

Period	Points	Setting of CRAC Return Temperature	(Measured Value-Based Run) Sum of 706 Points
21 November 2019–6 December 2019	706	30 °C	49.15 kW

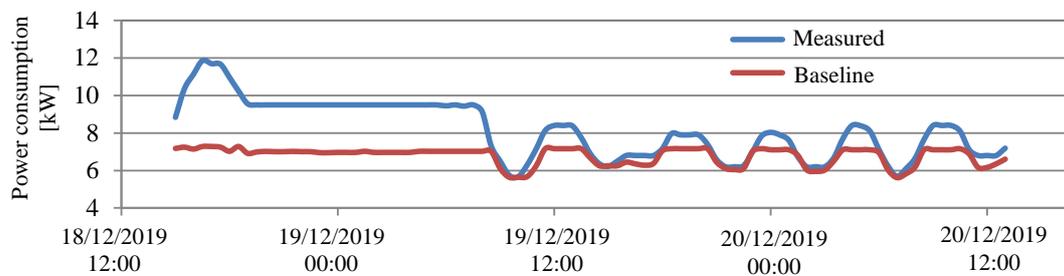


Figure 15. Comparison of measured power consumption and baseline time-series change when CRAC return temperature is changed to 30 °C.

Past data may be used to verify the energy-saving effect. To confirm the features of using the model using machine learning, the energy-saving effect was verified when power consumption in November 2019 was taken as the power consumption in the previous year (return-air temperature setting for CRAC was 28 °C). Measured values of power consumption for CRAC for each year and time-series changes of the values predicted by the baseline model using 2019 data are shown in Figure 16. As clear from this result, with the return-air temperature of the CRAC set at the same value (28 °C), and even if the periods (21 November to 30 November) are equivalent, it is difficult to treat the past record of CRAC power consumption as a baseline. This difficulty can be attributed to the difficulty in using last year’s data because internal load changes over time (as described above).

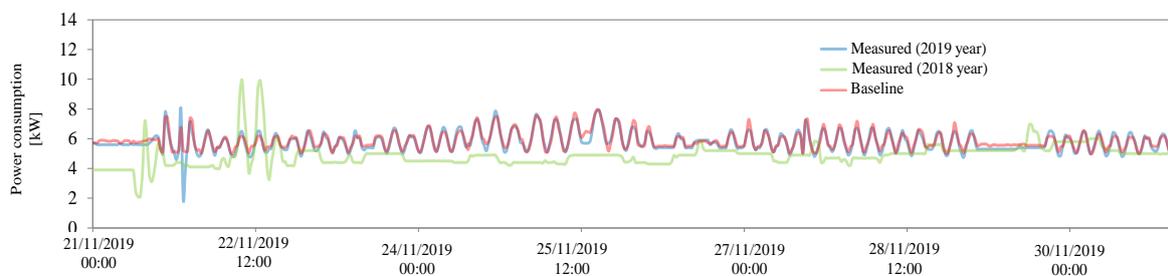


Figure 16. Comparison of actual power consumptions in 2018 and 2019 and baseline power consumption of CRAC.

5.4. Additional Verification of Baseline When the Setting of CRAC Return Temperature Is Changed

For the purpose to add explaining the certainty of the baseline when the setting of CRAC return temperature is changed as explained in the previous section, the verification when the setting of CRAC return temperature is lowered was conducted. The verification was carried out during the following period and the verification results are shown in Table 25. From this table, it is shown that when the setting of CRAC return temperature lowered, the actual power consumption of CRAC is higher than that of the baseline, and how much energy is increased when this change of CRAC return temperature is implemented. I think it is possible to quantitatively grasp what has happened.

Table 25. When return temperature of the CRAC is set to 20–26 °C (10 December 2019 to 20 December 2019; sample number: 479).

Period	Points	Setting of CRAC Return Temperature	(Measured Value-Based Run) Sum of 706 Points
10 December 2019– 12 December 2019	97	26 °C	−33.05kW
12 December 2019– 16 December 2019	193	24 °C	−37.76kW
16 December 2019– 18 December 2019	97	22 °C	−13.67kW
18 December 2019– 20 December 2019	92	20 °C	−126.97kW

6. Concluding Remarks

In this study, we proposed the methods which are prediction of rack intake temperature and calculation of baseline to support to promote energy-saving measures. The evaluation of indices was defined for each problem, the suitable machine learning method was considered for each problem, and the method was narrowed down by experiment and verification. Using the selected method, we constructed a model that can predict and calculate the rack intake temperature and baseline with high accuracy in a server room. Furthermore, we clarified the effects of model explanatory variables and learning data that affect data accumulation on accuracy, and added consideration on data collection and accumulation in the server room. By utilizing the proposed method and the data in the server room, it is possible to support a highly reliable and highly efficient data center operation. However, since this study is a result in one server room, research on the application of this technology in different rooms is an issue in the future.

- Prediction of rack intake temperature
 - We defined an evaluation indices that we considered to be important of data center operation, and verified it with multiple machine learning methods which has character of self-learning. I built a model which predicted using a state space model as a method high accuracy from the viewpoint of the evaluation indices.
 - It was clarified that the return temperature of CRAC is an important among the explanatory variables on this model.
- Calculation of the CRAC baseline
 - We selected a machine learning method with XAI that we thought was important in this problem.
 - We verified the multiple methods and selected GBDT as a method high accuracy from the viewpoint of evaluation indices. In addition, We quantified the influence of the explanatory variables on the objective variables and showed that the model has explanatory power.

Author Contributions: Both authors contributed to this work as follows: Conceptualization, K.S. and M.K.; Methodology, K.S.; Software, K.S. and T.A.; Validation, K.S. and M.K.; Formal Analysis, K.S.; Investigation, K.S. and M.K.; Resources, M.K.; Data Curation, K.S. and T.A.; Writing—Original Draft Preparation, K.S.; Writing—Review & Editing, K.S. and K.M.; Visualization, K.S.; Supervision, M.S. and T.W.; Project Administration, M.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest

References

1. Andrae, A.; Edler, T. On global electricity usage of communication technology: Trends to 2030. *Challenges* **2015**, *6*, 117–157. [CrossRef]
2. Eric, M.; Arman, S.; Nuoa, L.; Sarah, S.; Jonathan, K. Recalibrating global data center energy-use estimates, American Association for the Advancement of Science. *Science* **2020**, *367*, 984–986.
3. ASHRAE Technical Committee 9.9 (TC 9.9). *Datacom Equipment Power Trends and Cooling Application Second Edition*; American Society of Heating Refrigerating and Air-Conditioning Engineers Inc.: Atlanta, GA, USA, 2012.
4. Geng, H. *Data Center Handbook*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2015.
5. ASHRAE Technical Committee (TC) 9.9 Mission Critical Facilities, Data Centers, Technology Spaces, and Electronic Equipment. *Data Center Power Equipment Thermal Guidelines and Best Practices*; American Society of Heating Refrigerating and Air-Conditioning Engineers Inc.: Atlanta, GA, USA, 2016. Available online: https://tc0909.ashraetcs.org/documents/ASHRAE_TC0909_Power_White_Paper_22_June_2016_REVISSED.pdf (accessed on 3 August 2020).
6. Tsukimoto, H.; Udagawa, Y.; Yoshii, A.; Sekiguchi, K. Temperature-rise suppression techniques during commercial power outages in data centers. In Proceedings of the 2014 IEEE 36th International Telecommunications Energy Conference (INTELEC), Vancouver, BC, Canada, 28 September–2 October 2014.
7. Lin, M.; Shao, S.; Zhang, X.S.; Van Gilder, J.W.; Avelar, V.; Hu, X. Strategies for data center temperature control during a cooling system outage. *Energy Build.* **2014**, *73*, 146–152. [CrossRef]
8. Garday, D.; Housley, J. *Thermal Storage System Provides Emergency Data Center Cooling*; Intel Corporation: Santa Clara, CA, USA, 2007.
9. Tsuda, A.; Mino, Y.; Nishimura, S. Comparison of ICT equipment air-intake temperatures between cold aisle containment and hot aisle containment in datacenters. In Proceedings of the 2017 IEEE International Telecommunications Energy Conference (INTELEC), Broadbeach, QLD, Australia, 22–26 October 2017.
10. Niemann, J.; Brown, K.; Avelar, V. Impact of hot and cold aisle containment on data center temperature and efficiency. In *Schneider Electric's Data Center, White Paper*; Science Center: Foxboro, MA, USA, 2017.
11. Sakaino, H. Local and global dimensional CFD simulations and analyses to optimize server-fin design for improved energy efficiency in data centers. In Proceedings of the Fourteenth Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems, Orlando, FL, USA, 27–30 May 2014.
12. Winbron, E.; Ljung, A.; Lundström, T. Comparing performance metrics of partial aisle containments in hard floor and raised floor data centers using CFD. *Energies* **2019**, *12*, 1473. [CrossRef]
13. Lin, P.; Zhang, S.; Van Gilde, J. Data Center Temperature Rise during a Cooling System Outage. In *Schneider Electric's Data Center, White Paper*; Science Center: Foxboro, MA, USA, 2013.
14. Andrew, S.; Tomas, E. Thermal performance evaluation of a data center cooling system under fault conditions. *Energies* **2019**, *12*, 2996.
15. Thermal Energy System Specialists LLC—TRNSYS 17 Documentation, Mathematical Reference. Available online: <http://web.mit.edu/parmstr/Public/TRNSYS/04-MathematicalReference.pdf> (accessed on 17 June 2019).
16. Zavřel, V.; Barták, M.; Hensen, J.L.M. Simulation of data center cooling system in an emergency situation. *Future* **2014**, *1*, 2.
17. Kummert, M.; Dempster, W.; McLean, K. Thermal analysis of a data centre cooling system under fault conditions. In Proceedings of the Eleventh International IBPSA Conference, Glasgow, UK, 27–30 July 2009.
18. Efficiency Valuation Organization. *International Performance Measurement and Verification Protocol*; Efficiency Valuation Organization: Toronto, ON, Canada, 2012; Volume 1.

19. Jian, L.; Jakub, J.; Hailong, L.; Wen-Quan, T.; Yuanyuan, D.; Jinyue, Y. A new indicator for a fair comparison on the energy performance of data centers. *Appl. Energy* **2020**, *276*, 115497.
20. Maurizio, S.; Damiana, C.; Onorio, S.; Alessandra, D.A.; Alberto, Z. Carbon and water footprint of Energy saving options for the air conditioning of electric cabins at industrial sites. *Energies* **2019**, *12*, 3627.
21. Chao, D.; Nan, Z. Using residential and office building archetypes for energy efficiency building solutions in an urban scale: A China case study. *Energies* **2020**, *13*, 3210.
22. Cheng, J.C.; Das, M. A BIM-based web service framework for green building energy simulation and code checking. *J. Inf. Technol. Constr.* **2014**, *19*, 150–168.
23. Stefan, G.; Mohammad, H.; Jiying, L.; Jelena, S. Effect of urban neighborhoods on the performance of building cooling systems. *Build. Environ.* **2015**, *90*, 15–29.
24. Massimiliano, M.; Benedetto, N. Parametric performance analysis and energy model calibration workflow integration—A scalable approach for buildings. *Energies* **2020**, *13*, 621.
25. Demetriou, D.; Calder, A. Evolution of data center infrastructure management tools. *ASHRAE J.* **2019**, *61*, 52–58.
26. Brown, K.; Bouley, D. Classification of Data Center Infrastructure Management (DCIM) Tools. In *Schneider Electric's Data Center, White Paper*; Science Center: Foxboro, MA, USA, 2014.
27. Sasakura, K.; Aoki, T.; Watanabe, T. Temperature-rise suppression techniques during commercial power outages in data centers. In Proceedings of the 2017 IEEE International Telecommunications Energy Conference (INTELEC), Broadbeach, QLD, Australia, 22–26 October 2017.
28. Sasakura, K.; Aoki, T.; Watanabe, T. Study on the prediction models of temperature and energy by using dcim and machine learning to support optimal management of data center. In Proceedings of the ASHRAE Winter Conference 2019, Atlanta, GA, USA, 12–16 January 2019.
29. Matt, T. Gunning, Explainable Artificial Intelligence (XAI), Defense Advanced Research Projects Agency (DARPA), June 2018. Available online: <https://www.darpa.mil/program/explainable-artificial-intelligence> (accessed on 31 August 2020).
30. Amina, A.; Mohammed, B. Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access* **2018**, *6*, 52138–52160.
31. Germán, R.; Carlos, B. Validation of calibrated energy models: Common errors. *Energies* **2017**, *10*, 1587.
32. Huerto-Cardenas, H.E.; Leonforte, F.; Del, P.C.; Evola, G.; Costanzo, V. Validation of dynamic hygrothermal simulation models for historical buildings: State of the art, research challenges and recommendations. *Build. Environ.* **2020**, *180*, 107081. [[CrossRef](#)]
33. Harvey, A.; Koopman, S. Diagnostic checking of unobserved-components time series models. *J. Bus. Econ. Stat.* **1992**, *10*, 377–389.
34. Chen, T.; Guestrin, C. *XGBoost: A Scalable Tree Boosting Systems*; ACM: San Francisco, CA, USA, 2016.
35. Mehdi, M.; Mostafa, R.; Hamid, S. Temperature-aware power consumption modeling in Hyperscale cloud data centers. *Future Gener. Comput. Syst.* **2019**, *94*, 130–139.

