Optimization of Metro Central Air Conditioning Cold Source System Based on PCA-ANN Data Model

Ying Zhou[1], Xinmei Li[2]* and Dongfang Yang[2,3]

[1]Ningbo Rail Transit Group Co. Ltd. Construction Branch, Ningbo, China, [2]Nanjing Fuca Automation Technology Co., Ltd., Nanjing, China, [3]School of Energy and Power Engineering, Nanjing Institute of Technology, Nanjing, China

Due to the unique features of metro central air conditioning systems' architectural design and application scenarios, systems demand a greater degree of energy-savings than standard buildings. The central air conditioning system is the major energy user in metro stations, with the cooling source system accounting for a substantial portion. As a consequence, enhancing the energy efficiency of the cold source system is critical for optimizing the energy efficiency of the central air conditioning system. After analyzing the potential for energy-savings, we propose an energy-saving control technique for cold source systems based on the PCA-ANN data model. Firstly, an operating condition simulation was performed using operational data and cold source system equipment specifications. The effective operating data in the operational data-base was then filtered using the simulation data. Additionally, principal component analysis was used to examine the chosen dates. Finally, the fitted and calibrated data model was utilized to optimize the functioning of the cold source system. August's revised approach resulted in a 10.5 percent decrease in system energy consumption. In comparison to using non-optimized energy parameters, the suggested technique provides a variety of energy efficiency advantages.

Keywords: PCA-ANN, air conditioning system, energy saving, data modle, optimization

1 INTRODUCTION

The metro station serves as the primary location for staff distribution and equipment storage across the metro system. The major construction of the station is often located underground, where it is somewhat contained and immune to external weather variations. Metro stations operate in a dynamic environment, with a high passenger traffic and a high energy consumption of the air conditioning system. During metro operation, the air conditioning system consumes the majority of the station's energy, accounting for 40–50% of total power consumption (Feng and Lu, 2017). The air conditioning system at the metro station consumes a lot of energy owing to its lengthy operation. The energy consumption of station air conditioning systems is inextricably tied to the dynamic nature of the station's loads and the optimal management of the system (Liu, 2012).

The metro central air conditioning system is a typical nonlinear system with a high inertia, significant hysteresis, time variable behavior, and strong coupling. The refrigeration and cooling systems of the metro central air conditioning unit are intricately connected and subject to distinct limitations (Zhang, 2016). Additionally, the central air conditioner's working parameters vary dynamically in response to variations in load rate and outside climatic circumstances. As a result, the

real time dynamic control of the central air conditioning system is challenging to accomplish, as is system operation optimization.
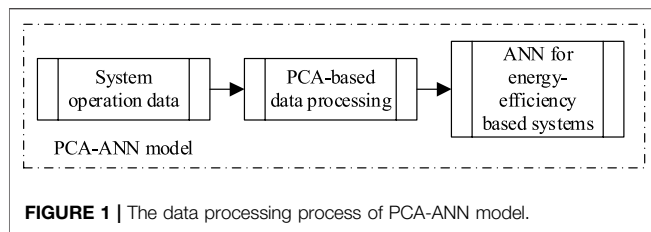
Literature (Wenzhuo et al., 2021) proposed a real time optimal control strategy adopting a multiagent based distributed optimization method for multi zone variable air volume (VAV) air conditioning systems. The proposed strategy consists of three novel schemes. The technique successfully balances thermal comfort, IAQ, and energy utilization while decreasing programming problems. Distributed optimization may provide nearly ideal results as centralized optimization. Literature (Yu et al., 2021) suggested a new iterative optimization algorithm based on a novel distributed control architecture and an alternating direction method of multipliers (ADMM) with regular term to address the problem of dynamic hydraulic balance regulation of chilled water pipe networks in central air conditioning systems. The suggested distributed optimization approach is highly accurate, stable, and efficient in terms of energy consumption. The proposed technology has the potential to save 28.54 percent of energy while preserving the pipe network's dynamic hydraulic equilibrium. The literature (Ning et al., 2017) developed a multi objective optimization methodology dubbed TOPSIS method for evaluating the task/ambient conditioning system's integrated performance and the optimal functioning was attained. According to the literature (Zhang et al., 2016), an energy consumption model for each component of the refrigeration system was presented, and the model was solved using methods such as surrogate lagrangian relaxation (SLR) and sequential quadratic programming (SQP). When compared to a baseline utilizing existing methodologies, numerical testing demonstrates that our approach results in considerable power consumption reduction, with the largest savings being roughly 18 percent in a partial load state at the maximum efficiency. The numerically modelled system's energy efficiency is higher under partial load (Diakaki et al., 2010). By examining an endless number of different metrics, the multi objective choice approach provided below seeks to circumvent this problem. These criteria are optimized using compromise programming, a well established multi objective solution paradigm. A simple case study is presented to show the decision model's functioning. The findings validate the approach's practicality, suggesting refinements and expansions. There exists in the scientific literature (Thangavelu et al., 2017) a technique for energy optimization that uses building thermal load and weather conditions to determine optimal operating choices for a chiller plant. Case study scenarios with various chiller plant layouts are used to evaluate the advantages of the suggested energy optimization technique. The results of the case studies revealed that energy-savings of up to 40% for medium sized chiller plants and about 20% for small chiller plants may be realized via better operations, resulting in lower energy costs and greenhouse gas emissions. The testing (Pisello et al., 2016) findings indicate that this technique of energy optimization is effective for medium sized central air conditioning equipment. This research reveals how historic buildings might create unfavorable interior environmental conditions for artwork preservation, inhabitant comfort, and building maintenance. This study proposes a numerical experimental approach to a

historic building. The study may help building operators and service managers evaluate and optimize interior environmental conditions for art preservation. Literature (Wang and Burnnet, 2001) established a simulation model for estimating a system's energy efficiency based on air conditioners physical properties. Online control optimizes the speed of variable speed condenser cooling water pumps. This allows you to optimize pressure management by changing your anticipated total power derivative. Adaptive control recognizes and adapts to changes in system characteristics. The literature (Meng et al., 2015) developed a control system that employed evolutionary algorithms to optimize air conditioner operating settings. Theoretically, all closed loop signals are constrained by Lyapunov synthesis.

However, mechanistic modeling of the circulation system can only restore the operating characteristics of some system components and cannot fully depict the relationship between each system component during operation; meanwhile, accurate modeling based on parameters also faces difficulties with convergence and operation speed during iterative calculation.

Numerous technologies, such as defect prediction and performance analysis, have benefited from the widespread use of deep learning and data mining methods. The operational data associated with central air conditioning systems are gaining increasing attention. Data mining methods are used to examine large amounts of equipment operating data and identify correlations between pieces of equipment as a foundation for system energy-savings. The present study is mostly focused on central air conditioning system diagnostics and is still in the research stage of energy-saving optimum operation.

The literature (Rasel et al., 2021) discusses the state of the art theory and application of the most widely used computational intelligence (CI) techniques, including the genetic algorithm (GA), particle swarm optimization (PSO), simulated annealing algorithm (SA), differential evolutionary algorithm (DE), chemical reaction optimization (CRO), multi objective genetic algorithm (MOGA), and non dominated sorting genetic algorithm II (NSGA II), for telecommunications. This holistic approach is well suited to the foresight of different forms of sustainable energy. Research (Daxin et al., 2021) talked about how to improve the automation level of a refrigeration plant and get the most out of energy-savings by using a group control system optimization strategy and a refrigeration plant model. This is what happened: The group control system could save 9.42 percent of its energy after it was optimized to run more efficiently. It could also save 5.67 percent after it was optimized to run more efficiently. An algorithm based on self adaptive differential evolution (SADE) is presented in literature (Dehao et al., 2021) as a global optimization technique for the vapor compression refrigeration system (VCRS) to decrease system energy consumption while meeting interior cooling needs. The global optimization problem is solved using the SADE method, and the optimum variable settings are found. By simulating the suggested global optimization technique, it has been shown that it is capable of increasing the VCRS energy efficiency over time while still meeting cooling needs. On a normal testing day, an

FIGURE 1 | The data processing process of PCA-ANN model.

FIGURE 2 | The central air conditioning system's simulation model structure.

average energy-savings of 15.57 percent may be realized. Additionally, partial cooling loads may save a lot of energy in the morning and at night. It is also shown that, in comparison to DE and PSO algorithms, the suggested SADE method can greatly cut computation time and prevent local minima, thereby proving its usefulness. This study provides a viable way for reducing air conditioning system energy use. According to the literature (Robert et al., 2020), the construction of a model for the refrigeration plant used to store mangoes that is capable of simulating both the chamber and the fruit temperatures is being considered. The model is capable of properly predicting the temperature profiles in the chamber, but it is more sensitive when modeling the temperature of the fruit. By using a fuzzy controller, you may attain the set point more precisely and in a shorter amount of time than you could with an on/off controller, while also saving energy in the process. According to the literature (Ning and Zaheeruddin, 2010), an optimal supervisory operation strategy based on neural networks was proposed in order to find the optimal set points for chilled water supply temperature, discharge air temperature, and VAV system fan static pressure so that the indoor environment can be maintained with the least amount of chiller and fan energy consumption. Simulated outcomes reveal that, as compared to the typical night reset operating scheme, the optimum operation strategy saves about 10% more power at maximum capacity under full load circumstances and 19% more power at reduced capacity under partial load situations. According to the literature (Chen and Chang, 2014), a neural network model of chiller energy consumption was developed, and a technique for finding the optimal chiller load distribution with the goal function of minimizing energy consumption was presented, which was based on the particle swarm algorithm. According to the literature (Le et al., 2019), the experimental data were compared to the results of the linear regression equal load allocation approach, which resulted in savings of 12.68 percent and 17.63 percent, respectively, as compared to the baseline. Literature (Li et al., 2017; Yabin et al., 2017) examined the energy consumption patterns of air conditioning system component loads using cluster analysis and association rule mining and discovered three unique energy consumption patterns. Fu and Fan. (2014) and Tronchoni et al. (2010) revealed significant operational patterns in historical building operation data by doing cluster analysis on historical building operation data.

It is difficult to determine the operational features of a refrigeration system via data analysis, and there is a great deal of uncertainty in dynamically constructing matched control methods based on the operating data. There is no equivalent study approach for assessing central air conditioning operational
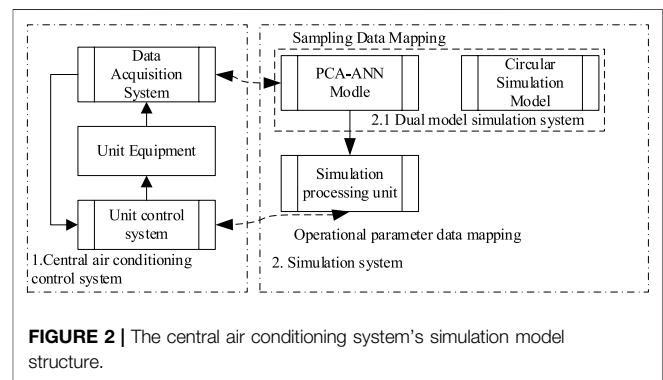
data due to the amount of time necessary to obtain and filter complete operational data. Data mining technology's effective acquisition, which is a vital component of the technology, will severely limit the unit's ability to work dynamically and optimally.

2 DATA COMPANION MODEL

Metro central air conditioning equipment is selected based on its ability to withstand the rigors of the work environment. Equipment operation and dynamic changes in external conditions cause the mechanism based equipment model to inaccurately depict the real operational characteristics and process changes of the equipment throughout the refrigeration cycle, resulting in equipment failure. The dynamic response of the system to its operational features based only on the gathered operating data, followed by the application of the appropriate control mechanism, resulting in a large number of uncertain data connections between the system components. The equipment settings for the central air energy-saving conditioning's optimization control plan, which is based on the simulation model, cannot be adjusted in order to account for the error caused by equipment performance loss in the simulation model. In addition, since the data-base based optimal control approach cannot entirely describe the system's operation al state, and because the data collection and accumulation time is inadequate, it is difficult to accurately depict the system's dynamic changes. To speed the optimization process for various combinations of central air conditioning systems, we recommend that we employ field operation and maintenance data, as well as system composition, to collect data from the field. It is discussed in the article how to implement an energy-saving control strategy for central air conditioning that is database based and integrates a dual-model verification operation approach that is based on both mechanism and data modeling. In order to achieve this, it is necessary to increase the energy efficiency of the central air conditioning system under all operating conditions, optimize the system's operating parameters, and forecast and optimize the simulation computation for the central air conditioning unit model, among other things. The operation of a central air conditioner accumulates a large amount of data, and a large data set with multiple variables will undoubtedly provide a wealth of information for research and application. In

many situations, the correlation that exists between variables may add complexity to the analysis of the problem. If each indicator is been analyzed separately, the analysis is often isolated and cannot fully utilize the information in the data, so blindly reducing the indicators can lose a lot of useful information and thus generate wrong conclusions. Therefore, it is necessary to propose a method to reduce the dimensionality of high dimensional feature data, keeping some of the most important features under the high dimensional data and removing the noise and unimportant features to achieve the purpose of improving the data processing speed.

PCA (Principal Component Analysis), the principal component analysis method, is one of the most widely used algorithms for data dimensionality reduction. The main idea of PCA is to map n dimensional features to k dimensions, which are brand new orthogonal features also called principal components, and are reconstructed k dimensional features based on the original n dimensional features. The work of PCA is to find a set of mutually orthogonal axes from the selection of new axes closely related to the data itself. This is equivalent to retaining only the dimension features containing most of the variance, while ignoring the feature dimensions containing almost 0 variance, so as to realize the dimensionality reduction processing of data features. Any change in form can be mathematically abstracted into a function. The mathematical expression of PCA was calculated using **Eq. 1**:

$$Z_{m\times k} = f(X_{m\times n}) \ k < n \qquad (1)$$

where X is the original data matrix and each row denotes an eigenvector of a sample. That is, matrix $X_{m \times n}$ contains m samples, each of which has n eigenvalues. As a result, PCA was employed to decrease the number of n.

The correlation distribution weights between the data are calculated during the correlation analysis. The data model's reference data will be validated by the features of the central air conditioning cycle. An artificial neural network (ANN) is a complicated network composed of a large number of basic components linked in a nonlinear fashion, capable of

performing complex logical operations and implementing systems with nonlinear relational realizations. When dealing with huge quantities of valid data, modeling using ANNs is an efficient approach. The following **Eq. 2** expresses the simple formula for ANN.

$$y = \sum softmax(\text{dot}(Wi, xi) + bi)) \qquad (2)$$

In the preceding equation, x signifies the input feature vector, and w denotes the weight assigned to each input feature, which indicates the feature's relevance. b signifies the influence threshold on the prediction outcome. y denotes the projected outcome. In the formula, the $dot()$ function represents the vector multiplication of w and x.

Although the operation of a central air conditioning cold source system conforms to the operating characteristics of a refrigeration cycle system, the operating characteristics of the system may be influenced by a variety of other operational parameters. The cold source system's operational characteristics may be modeled as a sophisticated continuous function. Theoretically, neural networks may approximate any preset continuous function with arbitrary precision. The neural network approach to central air conditioning operation enables the system to efficiently adapt to the huge quantity of data generated during operation. In general, the neural network methodology necessitates the use of statistical approaches to reduce the dimensionality. Data driven modeling techniques include the investigation of the unique patterns found in historical data throughout the mining process. Data driven methods are based on multivariate statistical techniques such as principle element analysis, principal regression, partial least squares, typical correlation analysis, Fischer discriminant analysis, and Hidden Markov Models. In our paper, we used PCA to reduce the data dimensionality of neural networks. As a result, we integrate PCA and ANN efficiently and propose a PCA-ANN based model for central air conditioning performance data.

In the actual application, we successfully integrated the data driven PCA-ANN model with the mechanism based simulation model to enhance the simulation process. The PCA-ANN
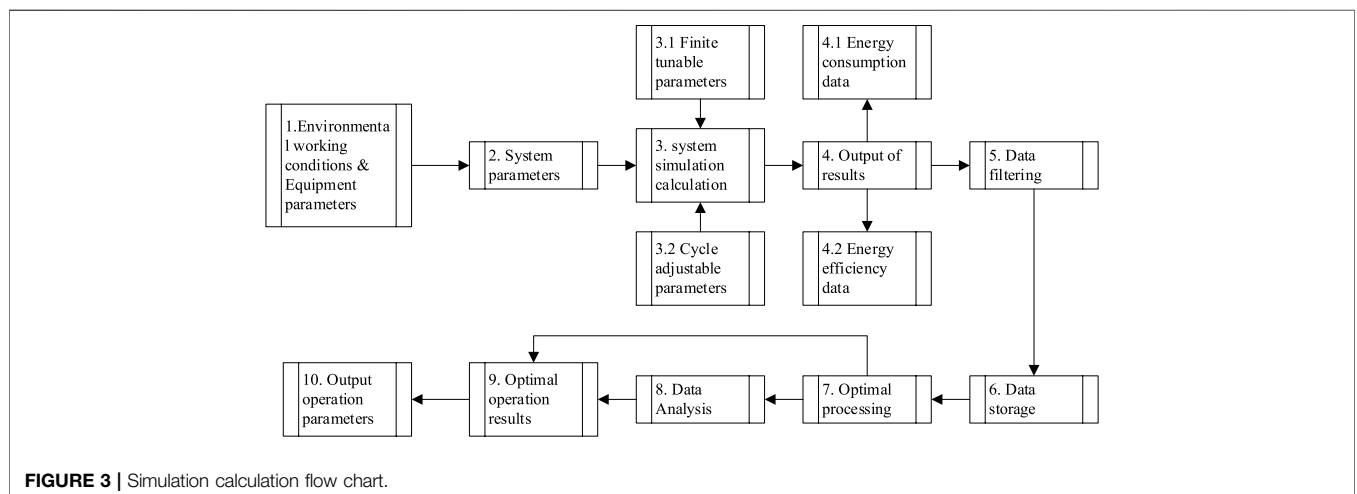

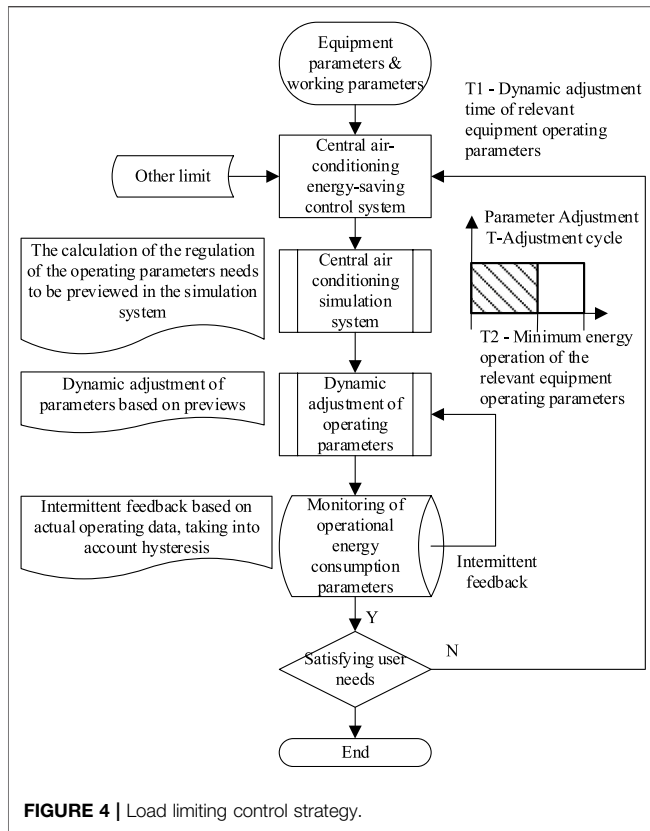**FIGURE 3 |** Simulation calculation flow chart.

FIGURE 4 | Load limiting control strategy.

model's data processing procedure is depicted in **Figure 1**. The central air conditioning system's simulation model structure is divided into two parts: the control system and the simulation system, with the simulation system consisting of the simulation-processing unit and the simulation system for the dual-model. The simulation-processing unit is utilized to combine the findings of the PCA-ANN and cyclic simulation models, which are subsequently used to control the system. **Figure 2** illustrates the simulation model framework of a central air conditioning system. **Figure 3** illustrates the unique physical simulation flow architecture for the data companion model of the central air conditioning system. The calculation flow is sequential and is divided into three major components: parameter input, system computation, and data analysis. The next section explains the refrigeration simulation system's simulation procedure for the portion of **Figure 3**.

3 DATA COMPANION MODEL OPERATION STEPS

The computation of the data companion model simulation is split into three stages: parameter specification, system calculation, data analysis and storage, with the following particular process steps. Step 1 is parameter input, Step 2 is parameter conversion, Step 3 is intermediate parameter calculation, Step 4 is power consumption calculation, Step 5 is calculation result output, Step 6 is data filtering, Step 7 is data storage, Step 8 is data comparison, Step 9 is

optimal operation results, and Step 10 is output operation parameters. Where Step2–Step7 is the procedure of calculating numerous cycles.

3.1 Step1–2 Parameter Definition

The parameter definition is divided into two sections: the definition of system external parameters and the definition of system rated parameters. Among them, the specification of system external parameters includes ambient dry bulb temperature and relative humidity, which can be collected in real time through the weather API, and the corresponding ambient wet bulb temperature may be calculated. Along with the unit's equipment parameters and external environmental parameters, the indoor dry and wet bulb temperatures and load rate are calculated as functions of the load limits, and in the event that other load limits are required, both the indoor environmental variables and the load rate must be dynamically adjusted to meet the demand, as illustrated in **Figure 4**. The time period for parameter adjustment in the figure is T, where $T = T_1 + T_2$, with the T1 time period devoted to dynamic adjustment of equipment related adjustable parameters and the T_2 time period devoted to equipment related adjustable parameters operating at the lowest possible power consumption.

3.2 Step3–6 System Calculation

The system calculation section is divided into two sections: the calculation of intermediate parameters and the computation of host parameters.

The system calculation section makes use of a method for modifying the parameter cycle computation. Heat exchangers are used to conduct heat exchange between water and refrigerant. In general, the same heat exchange, the greater the temperature differential between the heat exchange surfaces, the smaller the water flow, and the lower the pump's power consumption. To provide the most effective control approach, the temperature differential between the two sides of the heat exchanger is designed to fluctuate between (2–10°C) with a 0.1°C step. As with the cooling tower fan and pump, the cooling tower approximation is used to calculate heat transfer, with the approximation changing between (2–6°C) in 0.1°C increments. When calculating the temperature difference, it is essential to take into account both the pump's and cooling tower fan's frequency variations within an acceptable range (20–50 Hz), as well as the system's change in energy efficiency under a variety of configurations. Naturally, the calculation volume varies according on the operating conditions. In practice, all combination methods require that the cooling system have a specific amount of cooling capacity and be able to function, and throughout the cycle calculation process, some constraints on the range of combination methods for the pump host and cooling tower must be imposed. Additionally, keep in mind that the calculation of rationality may not match the real application requirements; thus, a single calculation on the water temperature must be back-calculated to see if it can meet the end of the need for chilled water temperature.

In the intermediate parameter calculation section, the calculation procedure must take into account not only the
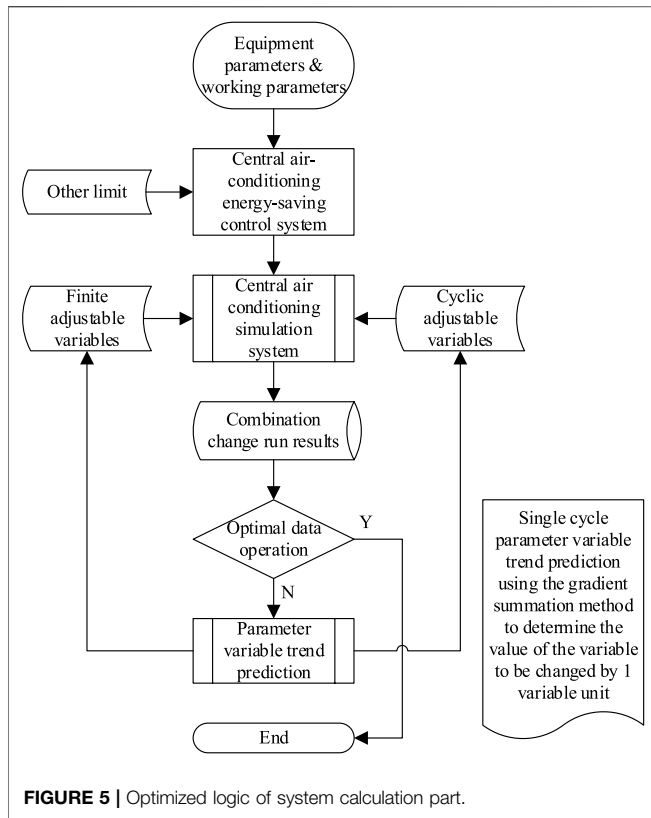
FIGURE 5 | Optimized logic of system calculation part.

number of pumps and their frequency of operation on the effect of water flow, but also the matching heat transfer efficiency of the heat exchanger's real form. At this stage, the weighting coefficients for heat exchange on both sides are included for a more precise computation.

The host parameters are calculated using the heat transfer correlation of the unit's performance, which is typically linearly fitted using the equipment's operating curve and actual observed data. The system heat exchange is estimated using the pressure enthalpy diagram's enthalpy difference as a function of the refrigeration cycle's parameter changes. However, modeling of the cycle computation needs knowledge of the equipment manufacturer's selection criteria and experimental data, as well as the installation of numerous measurement sites on the refrigeration system. Simultaneously, since the system is in the non-two-phase zone on the suction side, multiple back calculations are needed to obtain the enthalpy value used in the simulation calculation, resulting in a significant amount of computation. It is essential to compute the system heat exchange in combination with the actual operating needs and the equipment's operational characteristics during the simulation calculation procedure. Suction with liquid may efficiently decrease the exhaust temperature and improve energy efficiency in central air conditioning systems operating at high loads. The system computation section establishes a simulation model based on the system cycle using the equipment composition parameters. The model is primarily used to evaluate the features of the equipment at startup and to

forecast optimum performance; it is not directly engaged in the equipment's optimal operation control logic. At the start of the computation, the model is iterated repeatedly using finite and cyclic adjustable variables, and the optimum operation result under the operating circumstances is obtained using the gradient descent technique, as shown in **Figure 5**. Additionally, the load limit in the simulation computation has an effect on the calculation process of the host energy consumption. The upper limit of the data is computed in a single calculation based on the parameter adjustment time period, and the output energy consumption is calculated in conjunction with the prior dynamic regulation technique.

3.3 Step7–10 Data Analysis and Storage
The simulation results are saved in the database as arrays, and two primary activities, data comparison and data analysis, are conducted. The comparison of data is used to generate the operating strategy, while the analysis of data is utilized to improve the simulation computation.

3.3.1 Data Processing
After a specific quantity of data has been collected, it will be evaluated and computed. The control strategy and host energy consumption calculation parameters will be adjusted via analysis of the calculated and measured data. During the system simulation calculation process, a simulation model based on the system cycle is created using the characteristics of the equipment composition. The model is primarily used to evaluate the features of the equipment at startup and to forecast optimum operation; it is not directly engaged in the optimal operation of the equipment control logic. The model's calculation procedure is based on finite adjustable variables and cyclic adjustable variables for iterative computation; the gradient descent approach is used to find the
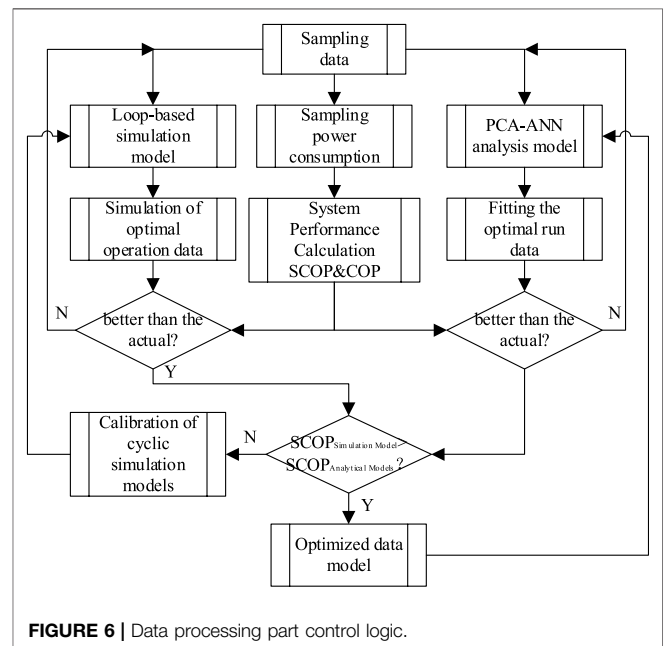

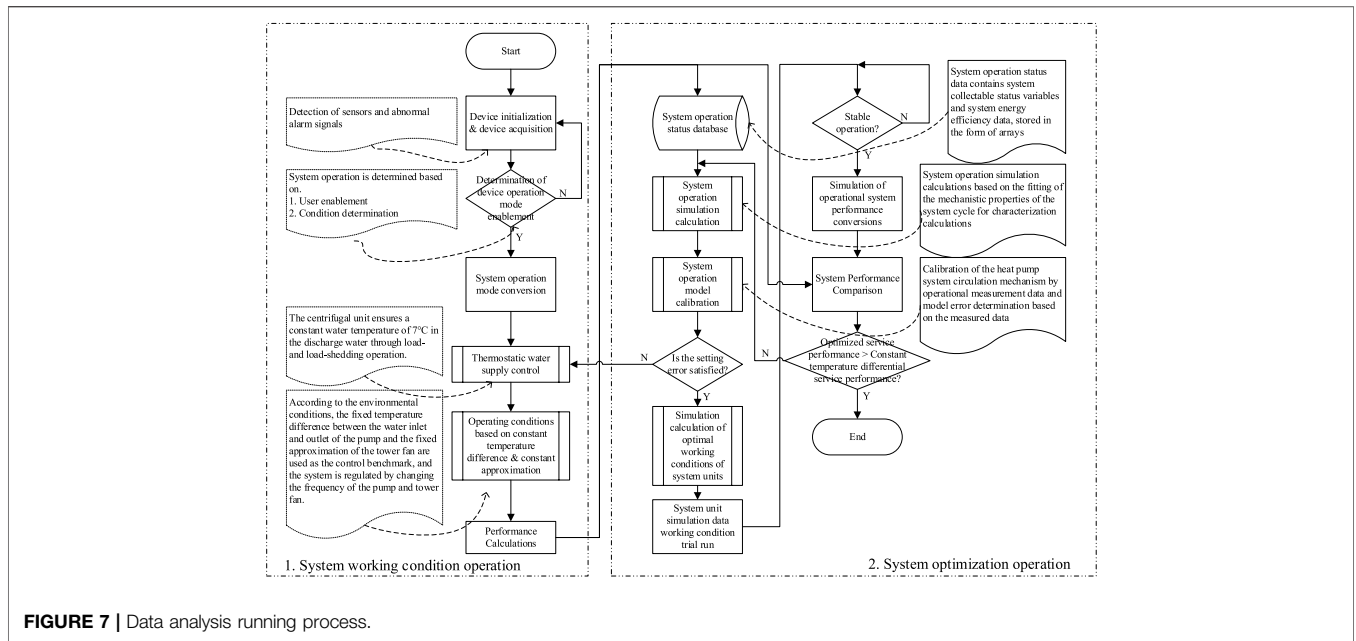**FIGURE 6 |** Data processing part control logic.

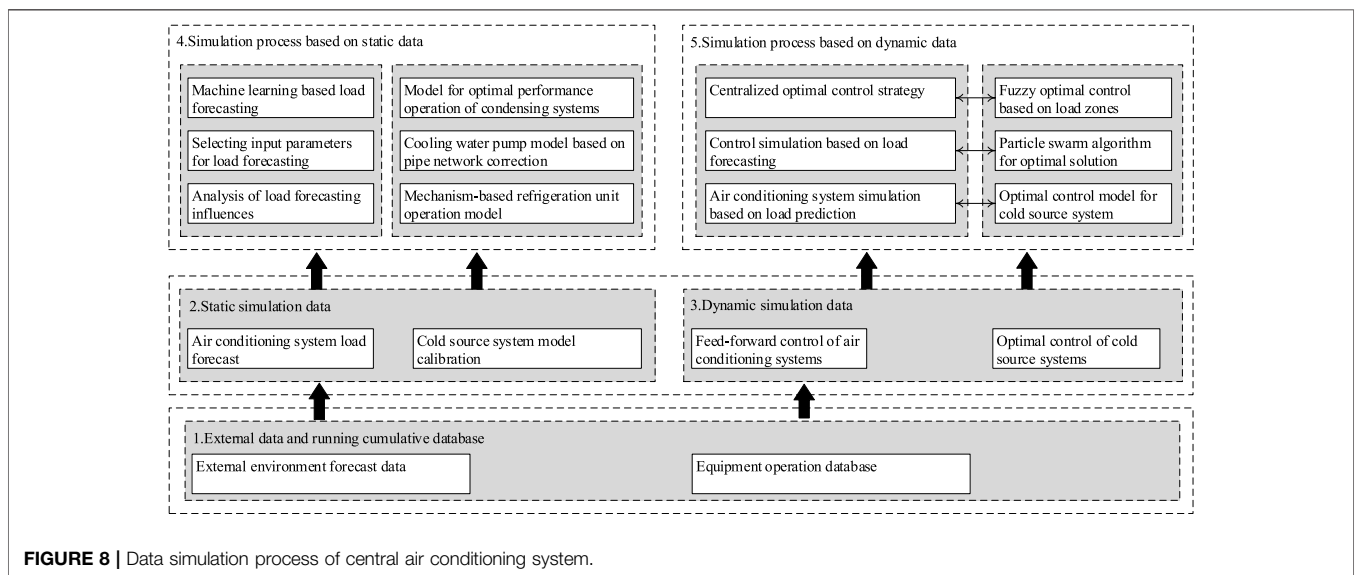**FIGURE 7 |** Data analysis running process.

**FIGURE 8 |** Data simulation process of central air conditioning system.

optimum operating outcomes for single working circumstances. Simultaneously, the load continues to have an effect on the computation of the mainframe's energy consumption. The data upper limit is restricted based on the time period of the parameters and coupled with the dynamic regulation technique to compute the output energy consumption in a single calculation.

The data analysis of the central air conditioner uses a dual-model design concept, as shown in **Figure 6**. At the start of the equipment's operation, the operation control strategy is mostly based on the simulation findings from the cycle-based simulation. Simultaneously, we revise the cycle model parameters in accordance with actual field measurement data, and we

evaluate the system's performance in relation to the COP and SCOP calculated by the system in order to make it more suitable for actual operation. The system's COP and SCOP calculation formulae are shown in **Eqs 3, 4**. The SCOP of a cooling system is the ratio of the cooling system's cooling capacity to its overall power consumption. The cooling system's overall power consumption is comprised of the refrigeration host, chilled water pump, cooling water pump, and cooling tower. The operation of the equipment creates a large amount of data on the system's operation. By comparing simulation and real measurement data, it is possible to quickly filter out and capture the optimal operation data in the form of an array including the operating conditions and operational parameters.

TABLE 1 | Central air conditioning system components.

Equipment name	Parameter	Quantity	Remarks
Centrifugal chiller 1	Rated power: 116.9 kW, cooling capacity: 813 kW	1	Power frequency
Centrifugal chiller 2	Rated power:138 kW, cooling capacity: 968 kW	1	Power frequency
Cooling Towers	Rated air volume: 126,400 m^3/h, flow: 208 m^3/h, power:4.4 kW	4	Frequency conversion
Refrigeration water pump 1	Rated flow: 110 m^3/h, power: 11 kW	2	Frequency conversion
Refrigeration water pump 2	Rated flow: 66 m^3/h, power: 7.5 kW	2	Frequency conversion
Cooling water pump 1	Rated flow: 176 m^3/h, power: 15 kW	2	Frequency conversion
Cooling water pump 2	Rated flow: 106 m^3/h, power: 7.5 kW	2	Frequency conversion
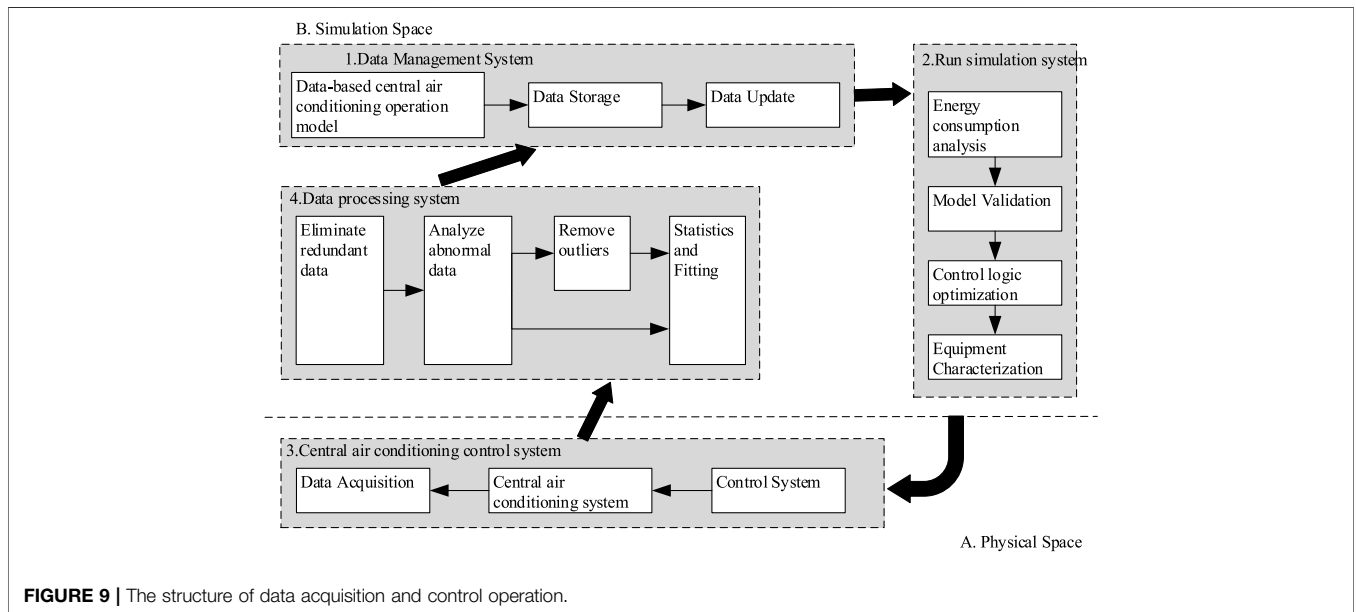
FIGURE 9 | The structure of data acquisition and control operation.

After a specific quantity of data has been accumulated, the system operating data will be processed efficiently. To begin, the data were analyzed using the PCA method in order to identify the connection between the optimum energy efficiency and the system operating parameter data, as well as the correlation distribution weights between the data. The data's significance was successfully filtered using the refrigeration cycle's rules and utilized as input variables for the data model. Second, we select an ANN neuron network to process the refrigeration system's model and create an efficient input-output model based on the system's characteristic characteristics and performance. The data model may be used to develop an efficient optimum operation control strategy for this goal. Simultaneously, the dual-model structure was refined and enhanced constantly *via* continuous examination of operating data.

$$COP = \frac{Q_{\text{Cold source system}}}{W_{\text{Refrigeration mainframe}}} \quad (3)$$

$$SCOP = \frac{Q_{\text{Cold source system}}}{W_{\text{Refrigeration mainframe}} + W_{\text{Cooling water pump}} + W_{\text{Refrigeration water pump}} + W_{\text{Cooling Towers}}} \quad (4)$$

The dual-model of the central air conditioning cooling system is intended to enhance the system's simulation computation. For optimal simulation results, tailored design is essential during model construction. The cold source system's mechanism model must be built using the equipment's characteristics, but the equipment's mechanism model requires several modifications based on real test results, and model accuracy is difficult to achieve in the early stage. To achieve more accurate simulation results and to guide real operation, it is often essential to run continuously for an extended length of time, and the dual-model's operating side rate is inefficient. However, such a design is essential in order to control the cold source system's whole life cycle.

3.3.2 Data Analysis

The data processing procedure necessitates data filtering based on the operating parameters of the system. As shown in **Figure 7**, valid data will be culled from the enormous data for further data processing.

Data analysis is the act of establishing a mathematical connection between the desired and controlled variables within each category. After calculating the refrigeration system's steady state performance, a restricted search for locally optimum solutions is performed using the gradient descent technique or an extremum search algorithm. From the
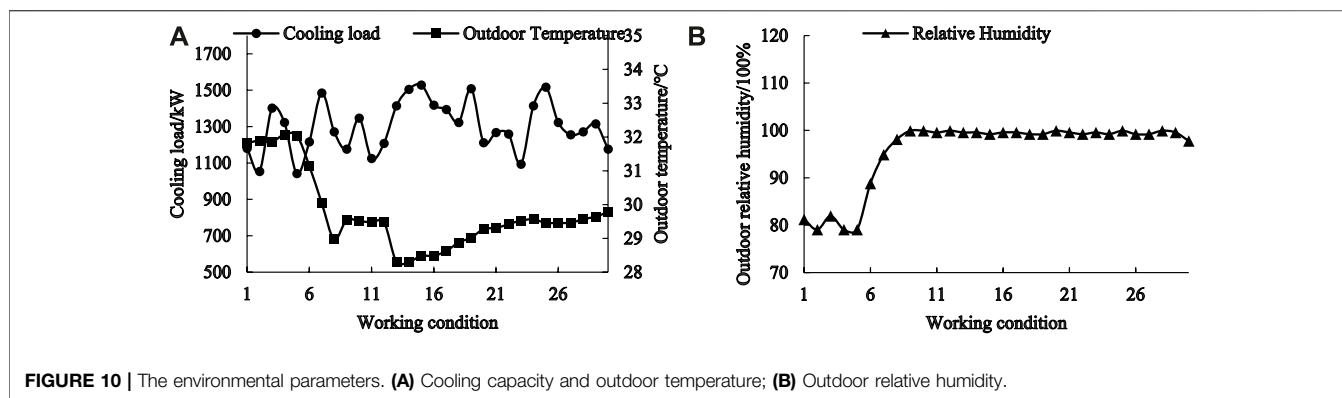
FIGURE 10 | The environmental parameters. **(A)** Cooling capacity and outdoor temperature; **(B)** Outdoor relative humidity.

TABLE 2 | Parameter setting of neural network.

Parameter	Value
Maximum number of training sessions	5,000
Number of implicit layer neurons	24
Network learning rate	0.00001
Training target error	0.00001

perspective of the system's steady state functioning, simulation calculations may effectively provide the system's correlation equation. Online correction of the simulation process is needed during the calculation process, based on the comparison of real operation data with simulation data under the equipment's operating circumstances. Following data analysis, operational dates are examined to create a database of system operating efficiency and system operating characteristics. When the operational circumstances replicate the database values or encounter an error, the system is controlled using the variable parameters specified in the data; when missing values occur, the system may be fitted using interpolation, with the values fitted using practical experience.

Following an analysis of the data processing process, the main stages involved in the process and implementation of the sampled data in the central air conditioning simulation system are detailed. The data simulation procedure for the central air conditioning system is shown in **Figure 8**. The central air conditioning simulation system is split into three levels: the database, static simulation data, and dynamic simulation data. External operation data and cumulative database of the central air conditioning system are divided into two sections: external environment prediction data and equipment operation data-base, which are used to generate alerting and dynamic simulation data, respectively.
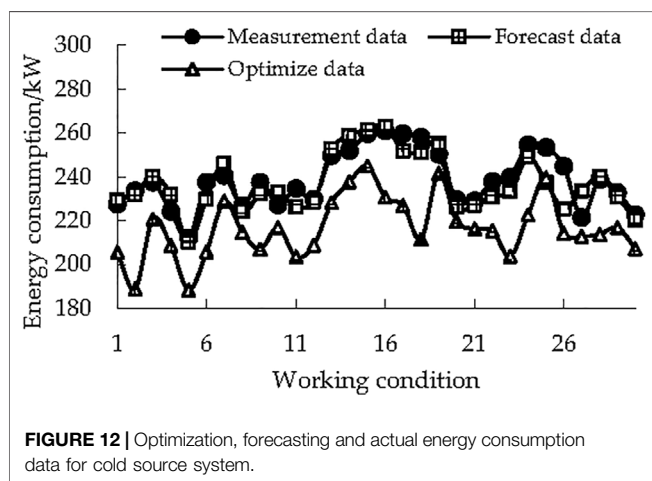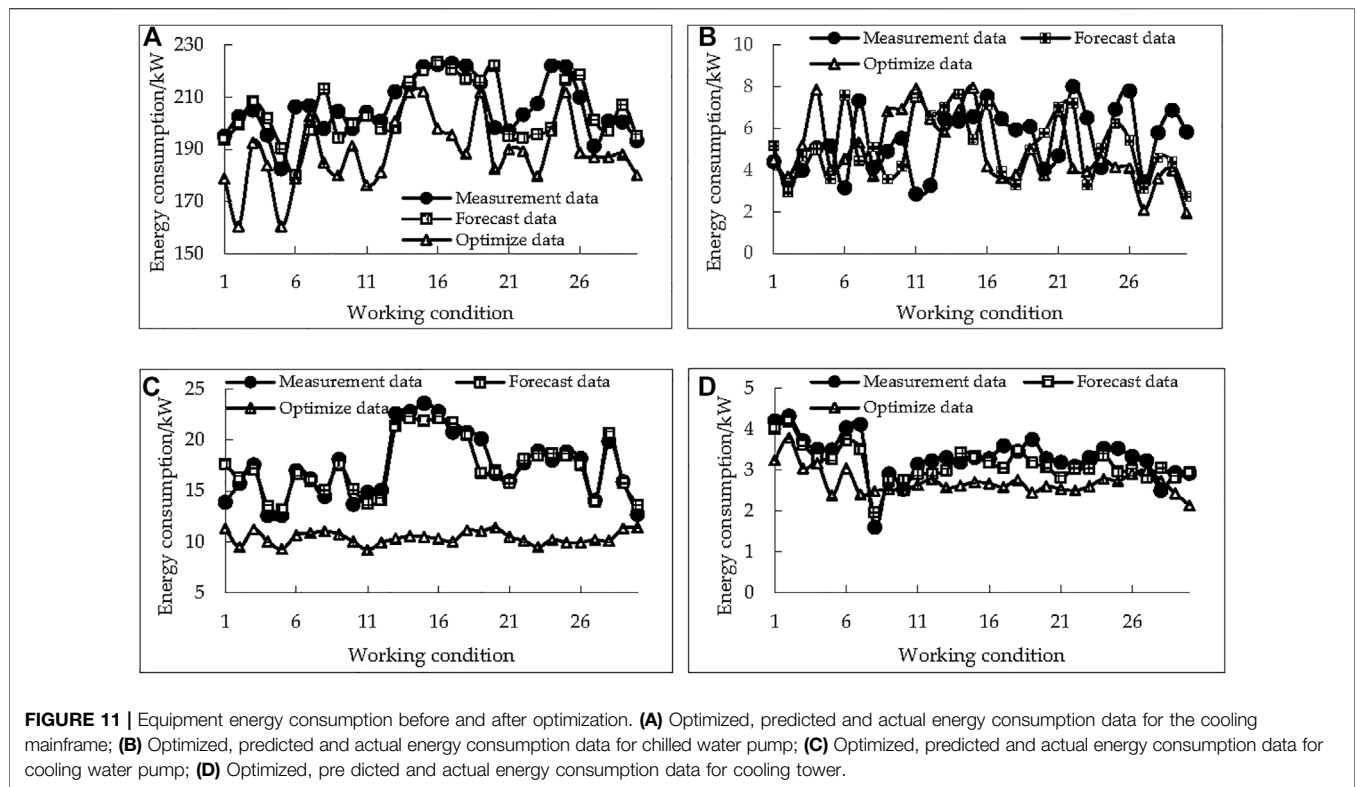
In the static simulation section, the air conditioning system load prediction forecasts the system's potential load situation at this moment based on external environment and operation data, while the cooling source system model performance parameters are calibrated using the equipment's operation data. The air conditioning system's compound prediction process comprises the computation of the compound prediction using machine learning, the selection of the input parameters for the compound

prediction, and the variables affecting the load prediction. Calibration of the cold source system models include establishing an optimum performance operation model, correcting the cooling water pump model using the corrected pipe network, and correcting the cooling unit operation model using the mechanism. In the dynamic simulation section, the simulation process of the dual-model offers advice on the air conditioning system's potential operating circumstances. The exchange of data and online transmission enables the cooling system to be controlled optimally. Air conditioning feedforward control consists of a centralized optimum control method, a control simulation based on load prediction, and an air conditioning system simulation based on load prediction. The optimum control of cold source systems related to feedforward control is implemented using fuzzy optimal control based on load zones, particle swarm optimal solution, and cold source system optimal control model, respectively.

4 DISCUSSION

To validate the aforementioned algorithm's efficacy, a collection of real world central air conditioning systems was simulated and operated; the system composition is given in **Table 1**. The simulation system has been put online, and **Figure 9** illustrates the structure of its data collection and control operation.

The online simulation system employs a two-level space structure, one for the physical environment and another for the simulation area. The physical space is defined as the central air conditioning control system, which consists primarily of data collection, central air conditioning, and control systems. The physical space is primarily responsible for the real functioning of the central air conditioning system and serves as the simulation system's execution component. The simulation space is comprised of three components: the data processing system, the data management system, and the operation simulation system. When the central air conditioning control system in the physical area receives operation data, it must first analyse it in order to accomplish redundant data rejection and abnormal data analysis and rejection. The processed data will next be statistically analysed

FIGURE 11 | Equipment energy consumption before and after optimization. **(A)** Optimized, predicted and actual energy consumption data for the cooling mainframe; **(B)** Optimized, predicted and actual energy consumption data for chilled water pump; **(C)** Optimized, predicted and actual energy consumption data for cooling water pump; **(D)** Optimized, pre dicted and actual energy consumption data for cooling tower.

FIGURE 12 | Optimization, forecasting and actual energy consumption data for cold source system.

and fitted to enhance the databased model of central air conditioning functioning. Finally, the data will be handled further by the data management system, which will include model refinement, data storage, and database updates. The valid data will be further analysed to determine the energy consumption of the system, to verify the system model, to improve the control logic, and to evaluate the equipment characteristics in the operation simulation system.

Thirty sets of operation parameters were optimized in this research during a typical summer cooling day at a Guangzhou Metro station with air conditioning equipment running from 7:00

to 23:00. The environmental characteristics associated with each operating state are shown in **Figure 10**, when the cooling load in the metro station is between 50 and 95% of its rated capacity and both chillers are operational.

The mathematical correlation equation for the central air conditioning operating load was designed based on the equipment selection parameters and the metro station's operating parameters. The set total parameter method was used to calculate the central air conditioning system's energy consumption under various operating conditions. The mechanism based simulation model's operating conditions are calculated using Python 3.8, and the system's operating energy efficiency is calculated using known environmental parameters and load rates. The PCA-ANN based data model is constructed using cumulative operational data from the metro central air conditioning system for data analysis, and the data processing is performed in the order depicted in **Figure 1**, in combination with the real data accumulation process. To begin, the accumulated operation data must be processed, which is mostly utilized to identify legitimate data sets within the operation data. After rejecting bad values, the screening process is based on the environmental working condition data and load rate as a known function, and the system energy consumption rate as a result of the target function for data screening, with the criteria for data screening being the data columns containing the optimal values of energy efficiency for the same working condition under the target function. The screened data are then processed using PCA, which can convert high dimensional variables with linear correlation into linearly uncorrelated low dimensional variables,

TABLE 3 | Central air conditioning system components energy-saving rate.

	The cooling Mainframe (%)	The chilled water pump (%)	The cooling water pump (%)	The cooling tower (%)	The cold source system (%)
Energy-saving rate	8.2	3.1	28.5	17.5	10.5

while retaining as many variables as possible from the original data. This allows the high dimensional data set to be mapped to the low dimensional space while retaining as many variables as possible. The PCA method makes use of a third-party Python 3.8 library named Sklearn. The PCA process is intended to lower the computing intensity of the future ANN and the computational time required to do real-time data processing. The reduced dimensional data set will be utilized to process the ANN model. To speed up the processing, we employ the standard BP-ANN approach for training the neuron model of acquisition variables running efficiency, and all computations are performed in Matlab 2018b's neuron module. After comparing the effect of parameter selection on model accuracy, we determined the ANN model structure to be 8-24-1, where 8 bits represent the input variables, 24 represents the number of hidden layers, and one represents the output variables. The specific implementation criteria are listed in **Table 2**.

The 30 sets of data collected during the actual procedure are analysed. The measured data are the energy consumption values derived from the working condition values, the projected data are the working condition values simulated using the mechanism based model, and the optimized data are the working condition values simulated using the du al model feedback system. All statistics on energy usage are shown in **Figure 11**. The comparison demonstrates a substantial reduction in the energy usage of the improved cooling water pump. However, the decrease in chilled water pump energy consumption is not apparent, since chilled water energy consumption is related to interior comfort, and user comfort is a critical factor in the simulation process. Although the cooling tower's energy consumption has been decreased, due to the cooling tower's modest energy consumption, it was not regarded as the primary energy-saving item during the simulation optimization. As the chiller system's primary energy consumer, it is essential to minimize its energy usage. As can be observed from the simulation results, when combined with the cooling system's energy control, its energy consumption decrease trend is more apparent. The chiller system's total performance is integrated, and we see a substantial decrease in data following simulation system optimization (**Figure 12**). By comparing data from 30 different workgroups, we can determine that the average energy-savings rate for the chilled water host is 8.2 percent, the average energy-savings rate for the chiller system is 10.5 percent, and the energy-savings rate for each component of the chiller system is as shown in **Table 3**.

5 CONCLUSION

The purpose of this article is to present an energy-saving control strategy based on PCA-ANN and mechanism simulation for the

purpose of optimizing the operation of central air conditioning systems. The approach offers a novel control concept based on dual-models calibrated against one another for energy efficient central air conditioning control. The strategy's viability is confirmed by simulation computations, and the full findings are presented in the article.

1) Model parameter correction procedures are incorporated into the central air conditioning system modeling process, and the air conditioning system's model variables are validated and corrected using operational data. Then the problems of performance degradation and matching deterioration that occur during air conditioning equipment operation will be satisfactorily resolved. Along with the physical model based simulation system, the system's simulation process incorporates a data-based operating model, which is calibrated and enhanced via a data exchange mechanism. A model structure of this type can greatly minimize the amount of time required to improve the data model. This system, however, necessitates frequent manual screening and maintenance. The accuracy of a dual-model simulation system is dependent on the availability of accurate data, and a certain quantity of data must be acquired before the simulation model may be optimized for equipment performance. However, the data collecting and filtering processes are frequently lengthy and imprecise due to factors such as sensor and operation. Even yet, in order to obtain reliable data for complete operating conditions, it is required to prepare the equipment prior to installation and commissioning, as well as to ensure that the entire central air conditioning system's components are capable of adapting to different operating conditions.

2) PCA approaches can be used to establish a correlation between operational data and the best performance of central air conditioning. Once a substantial amount of reliable data has been accumulated, the PCA method can be utilized to connect variable values for weight analysis. ANN networks can be successfully utilized to model data. The PCA algorithm and ANN approaches are coupled to build a PCA-ANN modeling scheme capable of successfully modeling a central air conditioning system with a small amount of data. To complete the simulation process using the currently utilized PCA-ANN model, interface communication via separate software is required, and the interface communication influences the computing rate due to the operating system's non real time nature. Improved integration of subsequent computational processes may result in an increase in computational efficiency.

3) We have developed the optimum sexual operation strategy in the face of limitations. The system has been guaranteed to

function optimally within defined limitations while maintaining stable refrigeration operating conditions. For this reason, the PWM operating strategy, in conjunction with the hysteresis characteristics of the refrigeration sys tem, may satisfy the user's requirements for periodic control. By integrating the operation al data from the metro stations, the system's energy-savings rate after simulation is 10.5 percent, which achieves the effective energy-savings of the air conditioning system while meeting the user's demand.

4) As an expanded idea, the dual-model research concept might be applied to different industrial domains. Many scholars have conducted previous studies for the simulation of mechanism models in various industries, and fruitful results have been obtained. Industry 4.0 has presented new challenges to many industries, and a lot of research has been conducted by many scholars in previous studies for the simulation of mechanism models in various industries, and fruitful results have been obtained. Data models have been applied to a wide range of practical applications as digital technology has progressed, and there are numerous successful examples. However, more research is needed to examine if the research subjects' research solutions can be transferred to industrial practical applications more quickly and successfully. It is possible to successfully combine current research results, and combining the outcomes of mechanism based model simulation and data-based computation will be an effective measure.

5) In our study, we discovered that the energy savings of central air conditioning cooling sources based on the PCA-ANN algorithm are mostly focused on the energy saving management of the cooling water pump and

cooling tower. Because the cooling side of the subway central air conditioning system normally uses the design demand of maximum load supply cooling, the energy saving impact on the cooling side is smaller than the effect on the heating side of the subway central air conditioning system. Current study is focused on typical metro load rates, which we are now performing. The findings of the simulation utilizing PCA-ANN, on the other hand, reveal that the cooling source system can save up to 10% of its energy consumption on average.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

Conceptualization, XL and YZ; Methodology, XL and YZ; Software, XL and YZ; Formal analysis, DY; Data curation, DY; Writing, DY.

ACKNOWLEDGMENTS

REFERENCES

Chen, C.-L., Chang, Y.-C., and Chan, T.-S. (2014). Applying Smart Models for Energy Saving in Optimal Chiller Loading. *Energy and Buildings* 68, 364–371. doi:10.1016/j.enbuild.2013.04.030

Daixin, T., Hongwei, X., Huijuan, Y., Hao, Y., and Wen, H. (2021). Optimization of Group Control Strategy and Analysis of Energy Saving in Refrigeration Plant. *Energ. Built Environ.* 2021 (2), 6. doi:10.1016/j.enbenv.2021.05.006

Dehao, K., Xiaohong, Y., Xudong, D., Ning, F., and Peiyong, D. (2021). Global Optimization of a Vapor Compression Refrigeration System With a Self Adaptive Differential Evolution Algorithm. *Appl. Therm. Eng.* 197. doi:10.1016/j.applthermaleng.2021.117427

Diakaki, C., Grigoroudis, E., Kabelis, N., Kolokotsa, D., Kalaitzakis, K., and Stavrakakis, G. (2010). A Multi-Objective Decision Model for the Improvement of Energy Efficiency in Buildings. *Energy* 35 (12), 5483–5496. doi:10.1016/j.energy.2010.05.012

Feng, A., and Lu, F. (2017). Statistics of Urban Rail Transit in Mainland China. *Tunnel construction* 38 (3), 514 517. http://www.suidaojs.com/CN/Y2018/V38/I3/514.

Guo, Y., Li, G., Chen, H., Hu, Y., Shen, L., Li, H., et al. (2017). Development of a Virtual Variable-Speed Compressor Power Sensor for Variable Refrigerant Flow Air Conditioning System. *Int. J. Refrigeration* 74, 73–85. doi:10.1016/j.ijrefrig.2016.09.025

Le, T., Lee, B., and Lee, G. (2019). The Development of the CS Valve Used to the Variable Swash Plate Compressor for a Vehicle Air Conditioning System. *Conf. Ser. Mater. Sci. Eng.* 604, 012–022. doi:10.1088/1757899X/604/1/012022

Li, G., Hu, Y., Chen, H., Li, H., Hu, M., Guo, Y., et al. (2017). Data Partitioning and Association Mining for Identifying VRF Energy Consumption Patterns under Various Part Loads and Refrigerant Charge Conditions. *Appl. Energ.* 185, 846–861. doi:10.1016/j.enbuild.2013.04.030

Liu, H. (2012). *Research on central Air Conditioning Dynamic Load Forecasting and Energy-Saving Optimal Control Mode. PhD Thesis.* Guangzhou, China: Guangzhou University.

Meng, W., Yang, Q., and Sun, Y. (2015). Adaptive Control of Variable-Speed Wind Energy Conversion Systems with Inaccurate Wind Speed Measurement. *Trans. Inst. Meas. Control.* 37 (1), 63–72. doi:10.1177/0142331214531008

Ning, M., Jingyu, H., Dongmei, P., Jing, D., and Mengjie, S. (2017). A Multi-Objective Study on the Operation of Task/ambient Air Conditioning Systems in Subtropics. *Energ. Proced.* 142, 1880–1887. doi:10.1016/j.egypro.2017.12.385

Ning, M., and Zaheeruddin, M. (2010). Neuro-optimal Operation of a Variable Air Volume HVAC&R System. *Appl. Therm. Eng.* 30 (5), 385–399. doi:10.1016/j.applthermaleng.2009.10.009

Pisello, A. L., Castaldo, V. L., and Pignatta, G. (2016). Integrated Numerical and Experimental Methodology for thermal Energy Analysis and Optimization of Heritage Museum Buildings. *Building Serv. Eng. Res. Technol.* 37 (3), 334 354. doi:10.1177/0143624415609910

Rasel, A., Shuhaimi, M., NorErniza, M. R., Kallol, B., Fahad, M., and Kamran, A. (2021). Artificial Intelligence Techniques in Refrigeration System Modelling and Optimization: A Multi Disciplinary Review. *Sustain. Energy Technol. Assess.* 47. doi:10.1016/j.seta.2021.101488

Robert, M., Manuel, M., Bogdan, D., Silupú, J. J. M., Alama, W. I., and Arellano-Garcia, H. (2020). A Fuzzy Control Approach for an Industrial Refrigeration

System. *Comput. Aided Chem. Eng.* 48, 1255–1260. doi:10.1016/B9780128233771.50210X

Thangavelu, S. R., Myat, A., and Khambadkone, A. (2017). Energy Optimization Methodology of Multi-Chiller Plant in Commercial Buildings. *Energy* 123, 64–76. doi:10.1016/j.energy.2017.01.116

Tronchoni, A. B., Pretto, C. O., and Rosa, M. (2010). Descoberta de conhecimento em base de dados de eventos de desligamentos de empresas de distribuio. *Saude e Sociedade* 21 (2), 7. doi:10.1590/s0103-17592010000200007

Wang, S., and Burnett, J. (2001). Online Adaptive Control for Optimizing Variable-Speed Pumps of Indirect Water-Cooled Chilling Systems. *Appl. Therm. Eng.* 21 (11), 1083–1103. doi:10.1016/S13594311(00)001095

Wenzhuo, L., Shengwei, W., and Choongwan, K. (2021). A Real Time Optimal Control Strategy for Multi Zone VAV Air Conditioning Systems Adopting a Multi Agent Based Distributed Optimization Method. *Appl. Energy* 287. doi:10.1016/j.apenergy.2021.116605

Xiao, F., and Fan, C. (2014). Data Mining in Building Automation System for Improving Building Operational Performance. *Energy and Buildings* 75, 109–118. doi:10.1016/j.enbuild.2014.02.005

Yu, J., Liu, Q., Zhao, A., Chen, S., Gao, Z., Wang, F., et al. (2021). A Distributed Optimization Algorithm for the Dynamic Hydraulic Balance of Chilled Water Pipe Network in Air-Conditioning System. *Energy* 223, 120059. doi:10.1016/j.energy.2021.120059

Zhang, D., Luh, P. B., Fan, J., and Gupta, S. (2015). A Decomposition and Coordination-Based Method for Chiller Plant Optimization,"in International Conference on Complex Systems Engineering. Dubai, United Arab Emirates: IEEE, 111 116. doi:10.1109/ComplexSys.2015.7385974

Zhang, Q. (2016). *Research on Energy-Saving Operation Control Method of central Air Conditioning systemPhD Thesis*. Nanjing, China: Southeast University.