ISSN: 2472-3703 DOI: 10.6911/WSRJ.202507_11(7).0011

Energy Consumption Prediction of Air-conditioning Based on SARIMA-GARCH Model

Chunli Kan[1, *], Qian Yu[1], Man Yang[1], Zhiyuan Hou[1] and Hongwu Yuan[1]

[1]School of Big Data and Artificial Intelligence, Anhui Xinhua University, Hefei Anhui, China

[*] Corresponding author

Abstract

Energy consumption prediction plays an important role in building design & retrofit, energy management system, but the nonlinear, dynamic and complex air-conditioning energy consumption data make it difficult. We employ SARIMA (Seasonal Autoregressive Integrated Moving Average) and GARCH (Generalized Autoregressive Conditional Heteroskedasticity) models respectively to extract the level and fluctuating information according to the trend, seasonality and random fluctuation of time series of energy consumption based on the real-time operation of air-conditioning in an office building, which are integrated to obtain a combined model of SARIMA(0,1,1)(1,1,1)$_9$-GARCH(1,2). The MAPE of SARIMA(0,1,1)(1,1,1)$_9$-GARCH(1,2) model is 4.46%.

Keywords

Air-conditioning energy consumption prediction; SARIMA-GARCH model; Seasonal time series; Heteroscedasticity.

1. INTRODUCTION

The total energy demand of a country is mainly that of the construction, industry, and transportation sectors. According to the World Watch Institute, building energy consumption is the largest component, accounting for about 40% of global energy consumption and about 36% of carbon emissions. With the growth of population, the shortage of total energy in the world has become increasingly serious, with immeasurable effects on economic development, the environment, and livelihoods. Therefore, it is a priority to save energy and improve energy efficiency. China has proposed "carbon peak" and "carbon neutral" goals. Air-conditioning constitutes an important part of building energy consumption. Accurate and reliable energy demand forecasts for air-conditioning can enable power companies to plan resources reasonably and balance supply and demand.

However, such a prediction is difficult, given the nonlinearity, dynamic nature, and complexity of air-conditioning energy consumption data, and the dependence of air-conditioning use on exogenous variables such as climate and social economy. Air-conditioning energy consumption prediction includes physical models and data-driven methods [1]. Data-driven energy consumption prediction is flexible, reproducible, and accurate and is widely practiced [2-3]. Most prediction models use multiple linear regression analysis and neural networks.

Although these methods can be used to predict air-conditioning energy consumption, the nonlinearity, trends, and seasonality of the data are ignored. In fact, air-conditioning energy consumption, which exhibits random fluctuation, is not only related to climate but is also affected by holidays and major emergencies. Time-series analysis is based on historical data, using curve fitting and parameter estimation to establish a mathematical model to predict future values, and is widely used in forecasting, such as [4]. Because of great inertia of building

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202507_11(7).0011

system, energy consumption at the next moment is closely related to its historical consumption, and seasonality and trend can be described by a seasonal model of time-series analysis, as [5]. In nonlinear time-series models, the variance of random disturbance terms is often unstable. To describe and predict this phenomenon, statisticians have proposed ARCH (Autoregressive Conditional Heteroskedasticity) and GARCH models, which complement classical ARIMA model. Ampountolas [6] evaluated the daily demand prediction performance of a GARCH model from multiple perspectives.

2. DATA SOURCE

The data used in this paper are real-time operation data of central air-conditioning in an office building, including energy consumption, load, and temperature of water at the inlet and outlet. The air-conditioner is turned off on holidays and during non-working hours. Therefore, the data from 8:00 a.m. to 5:00 p.m. on working days in June, July, and August of 2020 are adopted. The interval is 1 hour. For example, we regard the data of energy consumption at 8:00 as the data in the interval of 8:00 a.m. to 9:00 a.m., then there are 9 sets of data per day.

3. MODELING

3.1. Construct time series

The data is divided into two parts, data before August 16 for training, which constructs time series $\{x_t\}$. And data from August 16 to 30 for validation is to evaluate the prediction accuracy of the model. The 396 points of energy consumption data form the time series, shown as Figure. 1.

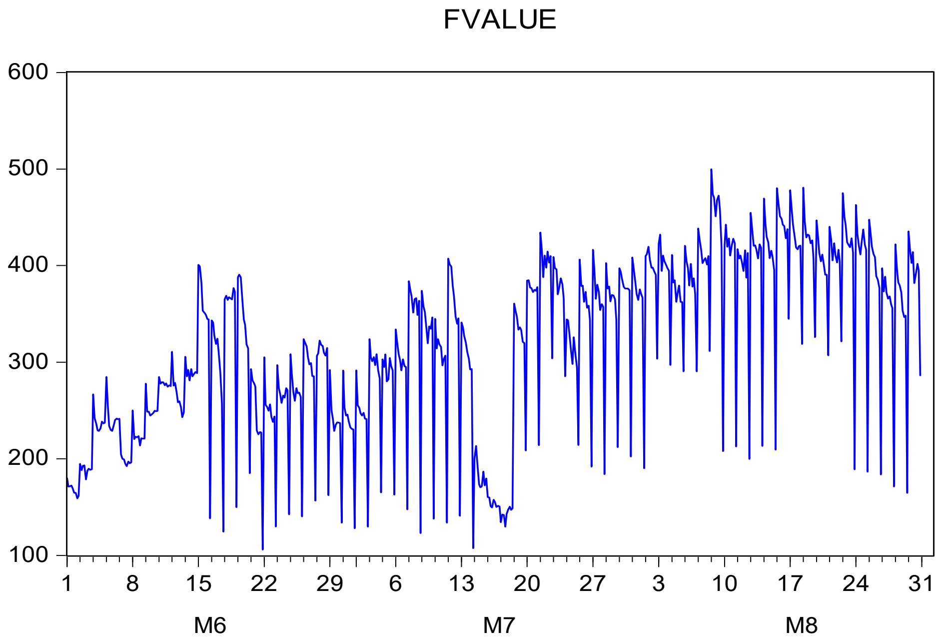


Figure 1. Time series of air-conditioning energy consumption

From Figure. 1, we can conclude that time series show seasonality and periodicity. Therefore, we can construct a seasonal time series model for it, and test stationarity and pure randomness.

3.2. Stationarity and pure randomness test of time series

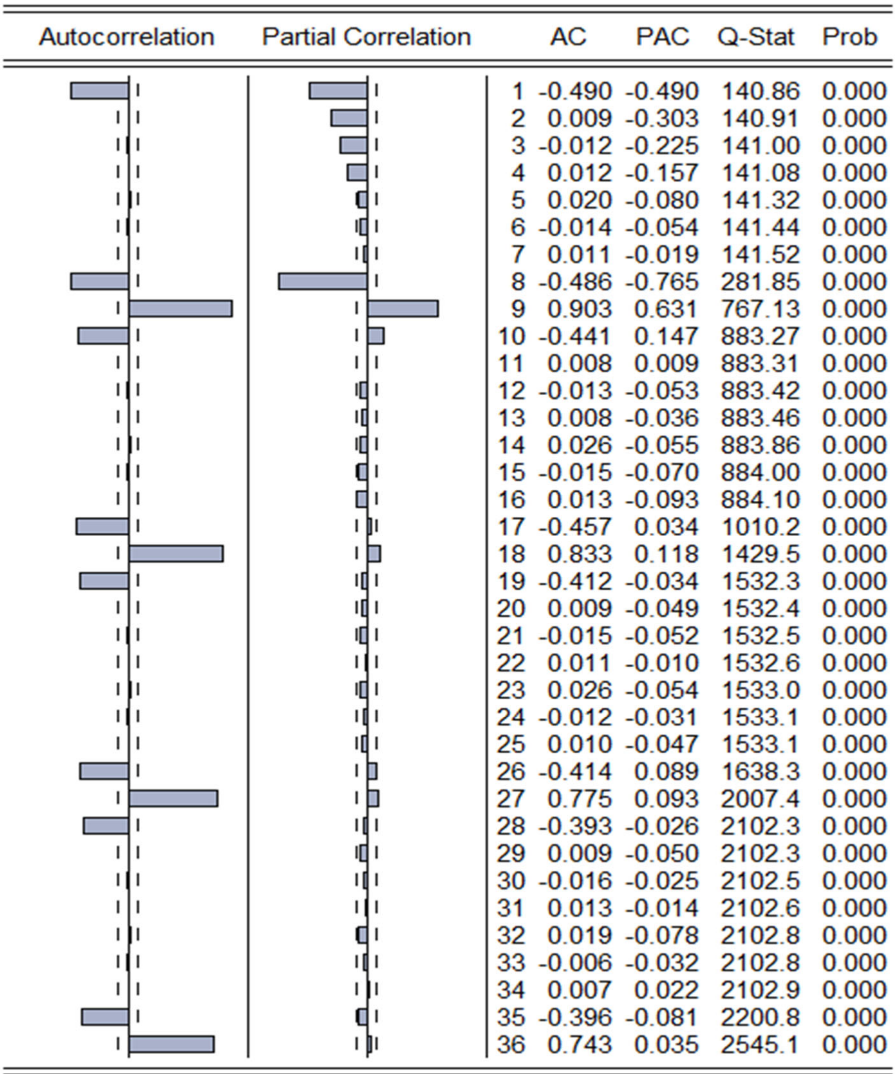
Before modeling, autocorrelation and stationarity tests are needed to determine whether time series method is suitable for energy consumption prediction. We conduct 1-order

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202507_11(7).0011

difference for the time series $\{x_t\}$ at first, the autocorrelations of time series after 1-order difference exhibits obvious seasonality and periodicity, and periodicity is 9, as shown in Figure. 2, then seasonal difference of order 9, also 9-step difference, is continued for the time series after 1-order difference. The results of a unit root test for the 1-order 9-step difference sequence is shown in Table 1. Because p-value is less than 0.0001, the null hypothesis is rejected, and we consider the difference sequence to be stable and no obvious periodicity exists in the series.

Table 1. Unit root test of 1-order 9-step difference sequence

t-Statistic	−31.62464
P	0.0000



	AC	PAC	Q-Stat	Prob
1	-0.490	-0.490	140.86	0.000
2	0.009	-0.303	140.91	0.000
3	-0.012	-0.225	141.00	0.000
4	0.012	-0.157	141.08	0.000
5	0.020	-0.080	141.32	0.000
6	-0.014	-0.054	141.44	0.000
7	0.011	-0.019	141.52	0.000
8	-0.486	-0.765	281.85	0.000
9	0.903	0.631	767.13	0.000
10	-0.441	0.147	883.27	0.000
11	0.008	0.009	883.31	0.000
12	-0.013	-0.053	883.42	0.000
13	0.008	-0.036	883.46	0.000
14	0.026	-0.055	883.86	0.000
15	-0.015	-0.070	884.00	0.000
16	0.013	-0.093	884.10	0.000
17	-0.457	0.034	1010.2	0.000
18	0.833	0.118	1429.5	0.000
19	-0.412	-0.034	1532.3	0.000
20	0.009	-0.049	1532.4	0.000
21	-0.015	-0.052	1532.5	0.000
22	0.011	-0.010	1532.6	0.000
23	0.026	-0.054	1533.0	0.000
24	-0.012	-0.031	1533.1	0.000
25	0.010	-0.047	1533.1	0.000
26	-0.414	0.089	1638.3	0.000
27	0.775	0.093	2007.4	0.000
28	-0.393	-0.026	2102.3	0.000
29	0.009	-0.050	2102.3	0.000
30	-0.016	-0.025	2102.5	0.000
31	0.013	-0.014	2102.6	0.000
32	0.019	-0.078	2102.8	0.000
33	-0.006	-0.032	2102.8	0.000
34	0.007	0.022	2102.9	0.000
35	-0.396	-0.081	2200.8	0.000
36	0.743	0.035	2545.1	0.000

Figure 2. Autocorrelations of time series after 1-order difference

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202507_11(7).0011

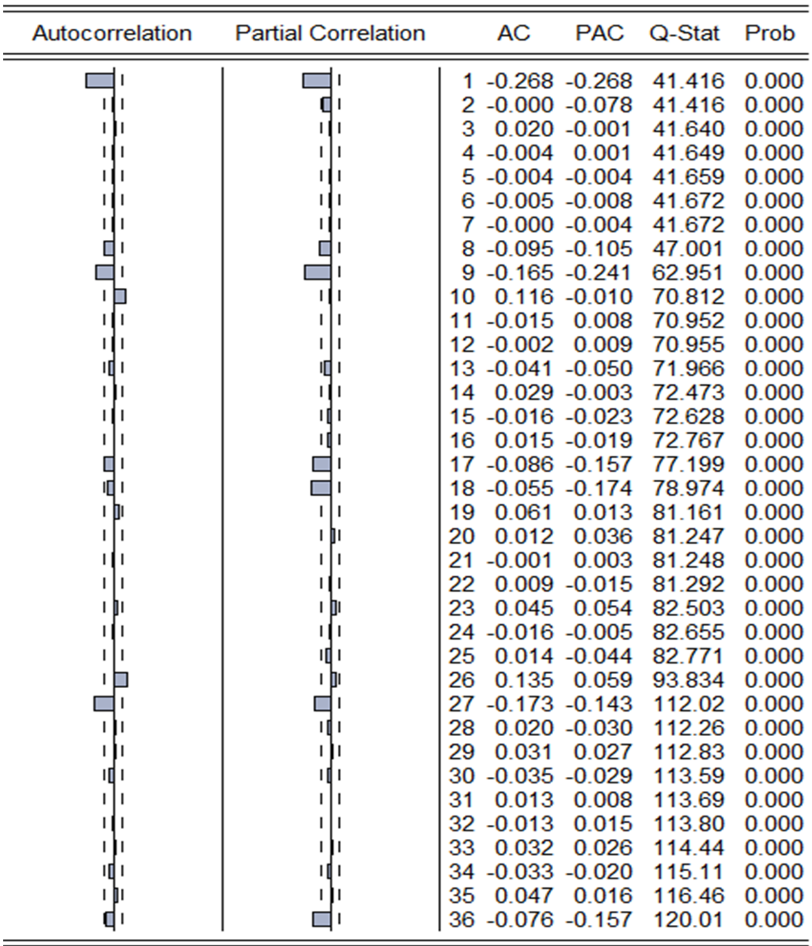


Figure 3. Autocorrelations of time series after 1-order 9-step difference

3.3. SARIMA model recognition

According to Fig. 3, the autocorrelation coefficients and partial autocorrelation coefficients of order 9 are trailing, so we employ ARIMA (1,1,1) to extract short-term autocorrelation information of series. The seasonal autocorrelation characteristics, trailing autocorrelation coefficients, and partial autocorrelation coefficients of delay of orders 9 and 18 are investigated. The seasonal autocorrelation information of the series was extracted by using SARIMA $(1,1,1)_9$ model with 9-step cycle. The model fitting results are shown in Table 2. Although the model passed the significance test, the constants C and AR (Autoregressive) failed. So, we need to refit the model.

Table 2. SARIMA $(1,1,1)\times(1,1,1)_9$ model fitting results

Variable	Coefficient	Standard error	t-Statistic	P-value
C	0.016495	0.092006	0.179282	0.8578
AR (1)	0.207576	0.106079	1.956818	0.0512
SAR (9)	0.453701	0.040326	11.25070	0.0000
MA (1)	-0.555128	0.089582	-6.196891	0.0000
SMA (9)	-0.958295	0.008988	-106.6196	0.0000

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202507_11(7).0011

Table 3. SARIMA $(0,1,1)\times(1,1,1)_9$ model fitting results

Variable	Coefficient	Standard error	t-Statistic	P-value
SAR (9)	0.452001	0.040392	11.19027	0.0000
MA (1)	-0.374881	0.040480	-9.260817	0.0000
SMA (9)	-0.959979	0.009116	-105.3123	0.0000

Table 3 shows the fitting results of SARIMA $(0,1,1) \times (1,1,1)_9$ model. The model as a whole passes the significance test, and each coefficient passes the test. If a fitting sequence passes the test, then the model can effectively fit the fluctuation of the observed value sequence at the specified significance level, but this is not the unique effective model. Therefore, the optimal model was determined by AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion), and HQ (Hannan-Quinn) information criteria.

Table 4. Minimum information scale

Model	AIC	BIC	HQ
SARIMA $(1,1,0)\times(1,1,0)_9$	9.620197	9.635549	9.626189
SARIMA $(1,1,0)\times(0,1,1)_9$	9.591302	9.606468	9.597217
SARIMA $(1,1,0)\times(1,1,1)_9$	9.444546	9.467574	9.453535
SARIMA $(0,1,1)\times(1,1,1)_9$	9.430514	9.453510	9.439489
SARIMA $(0,1,1)\times(1,1,0)_9$	9.609149	9.624480	9.615133
SARIMA $(0,1,1)\times(0,1,1)_9$	9.576908	9.592053	9.582815

The tested models are shown in Table 4. According to the minimum information criterion, the model that minimizes AIC, BIC and HQ functions is called the relative optimal model. So, the relative optimal model is determined as SARIMA $(0,1,1)\times(1,1,1)_9$.

3.4. SARIMA-GARCH combined model construction

The residual of SARIMA$(0,1,1)\times(1,1,1)_9$ model is shown as Figure. 4. It is observed that the fluctuation of residual series is stable in most periods, but it fluctuates greatly in some periods and keeps small fluctuation in some periods, so residual series need to be judged if they have heteroscedasticity.

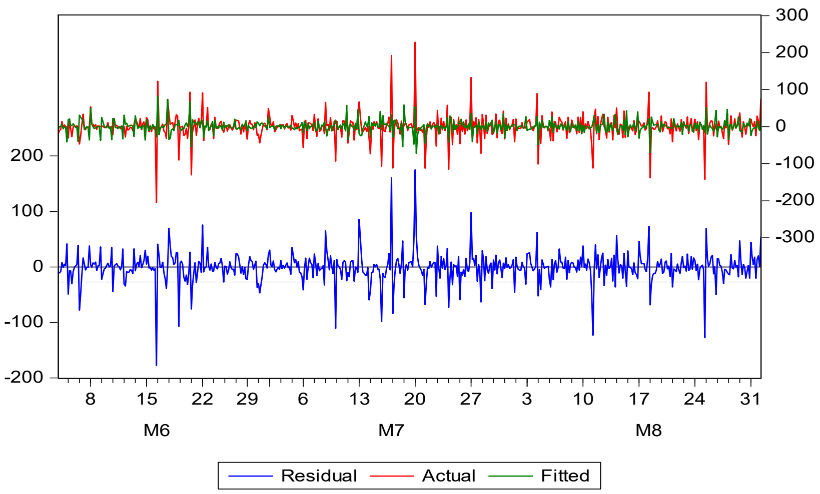


Figure 4. Residual

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202507_11(7).0011

In order to further judge whether residual sequence has heteroscedasticity, ARCH test is needed for residual sequence. The two commonly used ARCH test statistical methods are Portmanteau Q test and LM (Lagrange Multiplier) test, and the test results are shown in Table 5.

Table 5. ARCH test

Statistic	Value	P-value
F	7.572170	0.0061
nR^2	7.498221	0.0062

The p-values of Portmanteau Q test and LM test are 0.0061 and 0.0062 respectively, both less than the level of significance α=0.01. So, it is considered that the residual sequence is not homogeneous and has autocorrelation.

Table 6. SARIMA$(0,1,1)\times(1,1,1)_9$-GARCH (1,2) test

Variable	Coefficient	Standard error	t-Statistic	P-value
SAR (9)	0.531155	0.020520	25.88483	0.0000
MA (1)	-0.296509	0.022642	-13.09542	0.0000
SMA (9)	-0.962467	0.005339	-180.2866	0.0000
C	177.7435	38.23572	4.648626	0.0000
RESID (-1)^2	0.144891	0.026395	5.489355	0.0000
RESID (-2)^2	-0.135166	0.020695	-6.531421	0.0000
GARCH (-1)	0.741153	0.059374	12.48289	0.0000

Table 7. ARCH test

Statistic	Statistic value	P-value
F	0.100292	0.7516
nR^2	0.100631	0.7511

Table 6 shows the fitting effect of SARIMA$(0,1,1)\times(1,1,1)_9$-GARCH(1,2) model. It is observed that the coefficients of the model pass significance test. To ensure that the model is eliminated heteroscedasticity, ARCH test is conducted on the fitted model again. The P-values of the F and LM statistics are greater than 0.05 in Table 7. Therefore, we fail to reject the null hypothesis, and the constructed GARCH(1,2) model is considered to eliminate the influence of heteroscedasticity.

4. MODEL PREDICTION EFFECT

We take the energy consumption of air-conditioning in an office building as an example to establish a seasonal time series model. To evaluate the effect of the model, a cross-validation method is adopted. The prediction effect of SARIMA-GARCH model is based on MAPE,

$$\text{MAPE}=\frac{1}{n}\sum_{i=1}^{n}\frac{|y_i^{\mu}-y_i|}{y_i}\times 100\%,$$

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202507_11(7).0011

where y_i (1≤i≤n) is the actual value, y_i^{μ} is the predicted value, and n is the data size of the verification set.

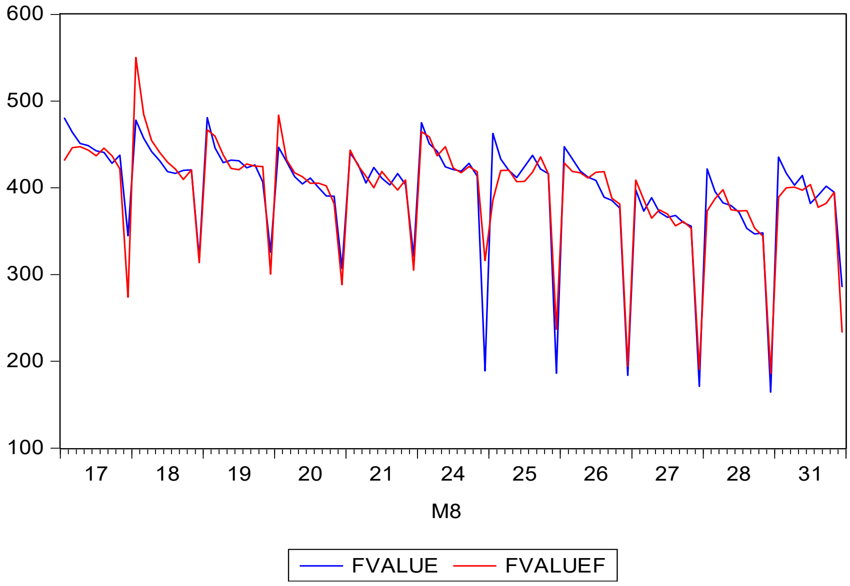


Figure 5. Prediction based on SARIMA-GARCH model

Table 8. Comparison of air-conditioning energy consumption prediction errors

Time	P	SARIMA-GARCH	absolute error
2020/8/17 8:00	480.3	432.0	48.3
2020/8/17 9:00	464.1	446.3	17.8
2020/8/17 10:00	451.2	447.4	3.8
2020/8/17 11:00	448.6	443.5	5.1
2020/8/17 12:00	442.5	436.8	5.7
2020/8/17 13:00	440.8	445.7	4.9
2020/8/17 14:00	428.4	437.0	8.6
2020/8/17 15:00	437.7	422.0	15.7
2020/8/17 16:00	344.9	274.0	70.9
2020/8/18 8:00	478.1	550.2	72.1
...	...	...	...
MAPE/%		4.46	

Table 8 presents the comparison of prediction effects of SARIMA(0,1,1)(1,1,1)$_9$-GARCH(1,2) model. p is energy consumption of air conditioner. The MAPE of SARIMA-GARCH model is 4.46%, and the prediction error is small.

5. CONCLUSION

We established a time series based on the real-time operating data of air-conditioning in an office building. The SARIMA model was used to extract the level information according to the trend, seasonality, and random fluctuation of the time series, and the GARCH model was used to extract the fluctuation information of the residual sequence. Then, the SARIMA(0,1,1)(1,1,1)$_9$-GARCH(1,2) model is obtained. Our conclusions are as follows:

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202507_11(7).0011

(1) The air-conditioning energy consumption data of the office building have obvious periodicity and seasonality. The series is stable after a nine-step difference, indicating that the period of energy consumption data is one day. For real life always makes air conditioning energy consumption exist random fluctuations for many different reasons, this article gives the accurate predictions of nine time points every day, makes the enterprise can more accurate grasp of the floating of air conditioning energy consumption changes, develop more controllable energy conservation and emissions reduction strategy.

(2) Because of the special nature of fixed commuting in office buildings, the energy consumption of air conditioning has obvious periodicity. Although the empirical model and VAR model can predict the energy consumption of air conditioning, they ignore the periodic change of energy consumption data. The SARIMA-GARCH model created in this paper predicts the future energy consumption in 9 time periods every day by depicting the periodic changes of historical data. This model is suitable for periodic data based on time dimension without considering variable factors.

ACKNOWLEDGMENTS

This paper was supported by The Excellent Scientific Research and Innovation Team Project of the Anhui Provincial Department of Education (2022AH010099), Anhui Province College Student Innovation Training Program (S202412216205), Anhui Xinhua University Intelligent Embedded Technology Research Center Project (yjs202109) and Anhui Xinhua University's school-level scientific research project (2024zr017).

REFERENCES

[1] J Liu, Y Zhang, K Wen: Analysis of different neural network models based on variational mode decomposition and dung beetle optimizer algorithm for the prediction of air-conditioning energy consumption in multifunctional complex large public buildings, Energy & Buildings, Vol. 334 (2025) No. 2025, p.115518-115518.

[2] B. R, C. M: Experimental investigation of influence of phase change materials in energy consumption of air-conditioning units and prediction performance evaluation of modified deep neural network model, Construction and Building Materials, Vol. 389 (2023) No. 4, p.123-132.

[3] X.Y. Li, S.Q. Chen, H.L Li: A behavior-orientated prediction method for short-term energy consumption of air-conditioning systems in buildings blocks, Energy, Vol. 263(2023) No. PD, p.45-61.

[4] S. Tayib, S. Nor, S.M. Norrulashikin: Forecasting on the crude palm oil production in Malaysia using SARIMA Model, Journal of Physics: Conference Series, Vol. 1988 (2021) No. 1, p.2334-2342.

[5] C. Zou, X.L. Liu: Urban Building Energy Consumption Forecast Based on the IPAT Theory, Advanced Materials Research, Vol. 2377 (2013) No. 4, p.482-486.

[6] Y. Zhang, Y.D. Peng, X.L. Qu: A Finite Mixture GARCH Approach with EM Algorithm for Energy Forecasting Applications, Energies, Vol. 14(2021) No. 9, p. 2352-2352.