

Deep Support Vector Machine-based Building Energy Consumption Prediction in a Data-Driven Model

Zhi-Min Lin¹, Hong-Li Zhu^{1,*}, Chee-Keong Tan²

¹School of Civil Engineering, Putian University, Putian 351100, P. R. China
13959505045@163.com zh1289798714@163.com

²Faculty of Engineering and Information Technology, INTI International University
71950 Jurai Newtown, Negeri Sembilan, Malaysia
ve3842@ucsiuniversity.edu.my

*Corresponding author: Hong-Li Zhu

Received July 10, 2024, revised December 9, 2024, accepted March 27, 2025.

ABSTRACT. *With global warming, energy saving and emission reduction has gradually become the consensus of various industries. Building energy consumption is a huge energy expense in China's construction field, and improving energy consumption prediction capability is a significant technical measure to enhance the level of fine management of building energy consumption. The data-driven energy consumption forecasting model has good prediction performance. Thus, this article designs a building energy consumption prediction method based on Deep Support Vector Machine (DSVM) in data-driven mode. Firstly, a multivariate linear regression model is adopted to analyze the intrinsic and extrinsic factors of the building energy system, and then a load model is constructed relied on the analysis results, and the inputs in the load model are taken as the influencing factors of energy consumption. Secondly, the historical building energy consumption data are processed with outliers, and then the influencing factor features are downsampled using Principal Component Analysis (PCA), and finally the downsampled feature vectors are divided into a number of DSVM segments, which are inputted into each sub-SVM for learning to obtain the deep-level features, and the feature vectors in each segment are assigned with the weights and reorganized into segments, so as to achieve the optimized SVM classifier, and the forecasting outcome of the energy consumption of the historical building are outputted. The experimental outcome implies that the suggested method has a small relative error in prediction, and the dispersion between the actual value and the predicted value of the energy consumption is also more desirable, so the prediction accuracy has a significant advantage.*

Keywords: Data-driven model; Deep support vector machine; Energy consumption prediction; Multivariate linear regression; Principal component analysis

1. **Introduction.** Recently, with the growth of the economy and the improvement of the living standard of the residents, the energy consumption of the building has been increasing. Accurate forecasting of building energy consumption is conducive to more reasonable energy planning, which is of great significance for strengthening the construction of energy monitoring system, promoting energy conservation and emission reduction in the field of building, and realizing the goal of carbon peaking as scheduled [1, 2]. Undoubtedly, the world is in the era of big data, the application of big data provides new opportunities for the development of the building energy efficiency industry, through the use of big data technology for in-depth intelligent analysis of building energy consumption and predictive

modelling can deepen the understanding of the building energy consumption and thus improve the level of energy management, and contribute to the early realization of China's carbon peak [3]. At present, research on building energy consumption has the following problems: first, for a long time, China's energy statistics are mainly aimed at the industrial aspect, and there is a problem of poor data quality in building energy consumption data [4]; second, although in recent years, China has already established a part of the energy consumption monitoring system [5], which has accumulated a large amount of data, most of the energy consumption data are not sorted out and made public. Therefore, how to improve the performance of energy consumption prediction in buildings under data-limited scenarios deserves further research.

1.1. Related work. With the increase in available data on building energy consumption, the field of building energy prediction has widely started to use data-driven models for prediction [6]. Data-driven models can be classified into two main categories: statistical models and machine learning models. Jahanshahi [7] used integrated moving average autoregression to predict building energy consumption and found that the ARIMA model was the best predictor. Nepal et al. [8] proposed a hybrid building energy prediction model combining clustering and ARIMA. Fumo and Biswas [9] developed a multivariate linear regression model to predict hourly and daily energy consumption of a research institute. Amber et al. [10] developed a multiple linear regression model to forecast the electricity consumption of a school building by using six environmental variables, and the results showed that the temperature and working day factors have a significant effect on the energy consumption of the building.

Due to the limitations of traditional data-driven prediction methods to deal with nonlinear problems, most researchers nowadays prefer to use machine learning methods for prediction. Support Vector Machine (SVM) is a widely used machine learning method, which is suitable for solving the problem of insufficient sample data and nonlinear problems. Mohandes [11] used SVM model to predict the energy consumption of buildings, and the study shows that the performance of SVM is better than autoregressive model. Dong [12] established SVM model to predict the energy consumption of commercial buildings, and the prediction results are more accurate and stable. Massana et al. [13] used a decision tree to predict the electrical energy consumption of buildings, and the performance of the prediction was low.

As the deep learning rapidly growing, Ekici and Aksoy [14] suggested a building energy consumption prediction model relied on Artificial Neural Networks (ANN). Zhang et al. [15] offered a PSO-RBFNN combinatorial model to predict the energy consumption of a building, but the prediction accuracy is not high. Li and Yin [?] suggested a building energy consumption prediction model relied on LM algorithm optimized BP network. Li and Su [16] developed a combined model of neural network and fuzzy inference system, and verified the effectiveness of the model in the prediction of energy consumption of buildings. Zhao et al. [17] developed an LSTM network model, which is used for the time-by-time energy consumption prediction of air conditioning systems. Khan et al. [18] constructed a CNN-based building energy consumption prediction model, which can accurately predict the value of building electricity consumption, but the prediction error is large. Yin et al. [19] analyzed the advantages and disadvantages of the three prediction models of BP neural network, support vector machine, and autoregression, and found that the prediction model based on SVM is more effective. DSVM is derived from traditional SVM, which maintains the advantages of SVM and the advantages of learning from neural network at the same time. Savi and Olivadese [20] proposed a deep learning algorithm based on SVM and applied it to the field of building energy prediction. Pallonetto et

al. [21] used the DSN architecture to organize the linear classifiers of the SVM for deep learning and used it for building energy prediction with good results.

1.2. Contribution. From the above analysis, it can be seen that the existing data-driven prediction methods have poor prediction efficiency. To enhance the prediction performance of building energy consumption, this article firstly adopts a multivariate linear regression model to analyze the intrinsic and extrinsic factors of the building energy use system, and constructs a load model based on the results of the analysis, and the inputs in the load model are taken as the influencing factors of energy consumption. Secondly, the historical building energy consumption data are preprocessed to eliminate abnormal data, and then the influencing factor features are downgraded using Principal Component Analysis (PCA) to remove data with low contribution rate. Finally, the downscaled feature vector is divided into many small DSVM segments, and multiple SVMs are used to learn different feature vector segments to obtain deep features, and the feature vectors in each segment are assigned weights and reorganized to the segments, to obtain the optimized SVM classifier, which outputs the prediction of historical building energy consumption.

2. Theoretical analysis.

2.1. Deep support vector machine. SVM is a single-layer learning machine, which maps the sample features in the original space to a new space by kernel function, and then constructs a linear discriminant function for classification. Multi-layer SVM is the basic theory of DSVM, and the new features established by mapping are used as inputs to train a new SVM to solve the classification problem [22].

DSVM has a total of k layers, the kernel function used in each layer is L_j , and s support vectors r_l ($0 < l \leq s$) are obtained by solving the SVM. The deep level features of a sample are determined by the inner product relationship between itself and all the feature vectors projected into the new space by the kernel function. Each element in the deep level feature $C = \{c_1, c_2, \dots, c_s\}$ ($0 < i \leq k, 0 < j \leq m$) is as follows.

$$c_{ij}^l = \beta_l y_l L_i(r_l, C_{(j-1)i}) \quad (1)$$

where β_l is the Lagrange multiplier, y_l is the label value corresponding to r_l , and $C_{(j-1)i}$ is the deep feature generated in the previous level. The feature dimension of each obscured level is the number of its support vectors.

Each node in the input level represents a corresponding element in the set $X = \{X_1, X_2, \dots, X_n\}$ of original sample feature vectors. The output level is the final SVM classifier trained with the k -th level of deep features, which is adopted to classify the unknown target with the following equation.

$$\begin{cases} F(x) = \sum_{l=1}^{s_f} \beta_l y_l L_f(r_l, D_k(x)) + a \\ y = \text{sign}(F(x)) \end{cases} \quad (2)$$

where s_f and L_f are the number of classifier feature vectors and kernel function, and $C_k(x)$ is the deep-level features extracted by the k -level DSVM network after training the target feature vector x to be classified.

2.2. Principal component analysis. PCA [23] can achieve dimensionality reduction by mapping the original high-dimensional eigen-set to a low-dimensional space through transformations. Assuming that the original sample is implied in Equation (3) and can be expressed as a linear combination of D variables as indicated in Equation (4).

$$X = [X_{(1)}, X_{(2)}, \dots, X_{(n)}]^T = [X_1, X_2, \dots, X_n] \tag{3}$$

where X_i is the feature vector and $X_{(i)}$ is the sample.

$$\begin{cases} Y_1 = a_{11}X_1 + a_{12}X_2 + \dots + a_{1D}X_D \\ Y_2 = a_{21}X_1 + a_{22}X_2 + \dots + a_{2D}X_D \\ \dots \\ Y_d = a_{d1}X_1 + a_{d2}X_2 + \dots + a_{dD}X_D \end{cases} \tag{4}$$

where Y_i denotes the i -th principal component, the principal components are uncorrelated and have the highest data variance in each direction.

Then the covariance matrix Σ is obtained and diagonalized. The sum of the contribution rates of the first m principal components with larger variance can be expressed as $\eta = \sum_{i=1}^m \lambda_i / \sum_{i=1}^D \lambda_i$. Then the feature vectors are extracted according to the contribution rates to complete the dimensionality reduction.

3. Analysis of factors affecting building energy systems. Real-time energy forecasting for buildings focuses on load analysis of energy-using systems as controllable loads. Building energy-using systems mainly include lighting systems, heating, ventilation and air-conditioning systems. In this paper, their loads are analyzed using a multivariate linear regression model [24], including the loads of the energy-using systems at that moment in time as affected by scattering, total radiation and outdoor temperature.

$$WZ_t = s_1Y_t + s_2TY_t + s_3TD_t + \varepsilon \tag{5}$$

where WZ_t describes the load generated by the energy-using system in the colour building at time t ; s_1, s_2 and s_3 represent the variables in the multivariate linear regression model; TY_t represents the total radiation generated by the energy-using system at time t ; TD_t represents the corresponding scattering value of the energy-using system at time t ; and ε represents the error term in the linear regression equation.

The energy system consists of two typical manoeuvres, the off-compressor manoeuvre and the on-compressor manoeuvre, which are indicated by the numbers 0 and 1, respectively. Denote by w_{air} the specific heat capacity of air and let $Q^{(i)}(t)$ represent the heat exchange that exists between the outdoor air and the building's energy system, which can be obtained from Equation (6).

$$Q^{(i)}(t) = \begin{cases} v_{air} q_{flow}^{(i)} [Y_{out}^{(i)}(t) - Y_{room}^{(i)}(t)], & control^{(i)}(t) = 1 \\ 0, & control^{(i)}(t) = 0 \end{cases} \tag{6}$$

where $Y_{out}^{(i)}(t)$ describes the actual outdoor temperature at time t ; $Y_{room}^{(i)}(t)$ describes the actual indoor temperature at time t ; $control^{(i)}(t)$ represents the control variable of the building's energy system at time t ; $q_{flow}^{(i)}$ describes the total mass of air heated or cooled by the energy system in the building per unit of time, which can be calculated by the following equation.

$$q_{flow}^{(i)} = \rho_{air} Flow^{(i)} S_t^{(i)} = \rho_{air} B_{air}^{(i)} \tag{7}$$

where ρ_{air} describes the air density, $Flow^{(i)}$ describes the ventilation of building i , $S_t^{(i)}$ represents the area corresponding to the vents, and $B^{(i)}$ represents the volume of air flowing through the building's vents per unit time.

Let $Q^{(i)}(t)$ represent the amount of cooling/heating generated in the building by the energy system per unit of time, as implied in Equation (8).

$$Q^{(i)}(t) = \begin{cases} v_{air} q_{heat}^{(i)} [Y_H^{(i)}(t) - Y_{room}^{(i)}(t)], & control^{(i)}(t) = 1 \\ 0, & control^{(i)}(t) = 0 \end{cases} \quad (8)$$

where $Y_H^{(i)}(t)$ represents the outlet temperature of the energy system in the building.

Based on the results of the above calculations, the electrical power $P^{(i)}$ of the energy-using system in the building is obtained, and its expression is as bellow.

$$P^{(i)}(t) = \frac{Q^{(i)}(t)}{EER^{(i)}} \quad (9)$$

where $EER^{(i)}$ stands for energy efficiency ratio.

Based on the results of the calculation of the electric power $P^{(i)}$, the electricity consumption $C^{(i)}$ of the energy-using system in the green building during T hours is calculated.

$$C^{(i)} = \int_0^H P^{(i)}(t)c^{(i)}dt \quad (10)$$

where $c^{(i)}$ represents the price of electricity.

The effect of the current moment load on the historical load is the internal factor.

$$WZ_t = n_1WZ_{t-1} + n_2WZ_{t-2} + \dots + n_nWZ_{t-n} + n_{n+1} \quad (11)$$

where n_1 , n_2 , and n_n represent historical load factors; n_{n+1} describes the error term; WZ_{t-1} represents the load generated by the energy system at the time $t-1$; WZ_{t-2} represents the load generated by the energy system at the time $t-2$; and WZ_{t-n} represents the load generated by the building energy system at the time $t-n$.

Thus, the load analysis model of the building energy system is constructed by combining the external factor model and the internal factor model.

$$VZ_t = x_1Y_t + x_2TY_t + x_3TD_t + x_4WZ_{t-1} + \dots + x_nWZ_{t-n} + \varepsilon \quad (12)$$

where ε describes the error term; x_1, x_2, \dots, x_n both represent the input variables, i.e., the factors influencing the energy consumption system.

4. Deep support vector machine-based building energy consumption prediction in a data-driven model.

4.1. Pre-processing of building energy consumption data. Intending to the issue of poor prediction effect of existing building energy consumption prediction methods, this article suggests a building energy consumption prediction method based on DSVM in data-driven mode. Firstly, outlier processing is implemented on historical building energy consumption data, and then PCA is used to downscale the features of the influencing factors of energy consumption and remove feature redundancy, and finally, the feature vectors of the DSVM segmentation are input into all the sub-SVMs for learning, and the feature vectors in the pairwise segments are reorganized for segmentation, so as to obtain the optimized SVM classifier, and the outputs are obtained as the prediction results of the historical building energy consumption. The entire model of the designed method is implied in Figure 1.

Before building energy consumption prediction, it is necessary to pre-process the historical energy consumption data to ensure its completeness and accuracy. Input parameters such as outdoor meteorological parameters are highly uncertain and complex. Random

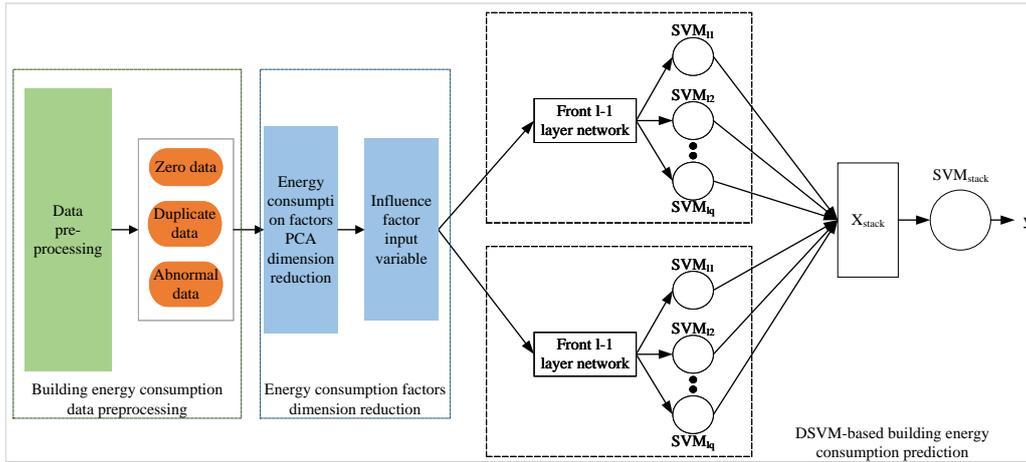


Figure 1. The entire model of the designed method

interferences such as system and equipment failures, as well as human errors in the operation of the energy-use system can easily generate erroneous historical energy consumption records, so abnormal data should be eliminated.

- (1) Enter the historical energy consumption data set, denoted $A = \{a_1, a_2, \dots, a_n\}$.
- (2) A data a_i is randomly selected in $A = \{a_1, a_2, \dots, a_n\}$.
- (3) Calculate the Euclidean distance between a_i and the rest of the data in dataset $A = \{a_1, a_2, \dots, a_n\}$, denoted as d_{ij} , using the following equation.

$$d_{ij} = \sqrt{\sum_{a_i, a_j \in A} (a_i - a_j)^2} \tag{13}$$

- (4) Sort d_{ij} from smallest to largest and select the top k nearest neighbours close to a_i , denoted as $b_l, l = 1, 2, \dots, K$.
- (5) Calculate the degree of isolation $B_i = \sum_{l=1}^k d_{il}$ of a_i , where d_{il} represents the sum of the Euclidean distances between a_i and the k nearest neighbour points.
- (6) Calculate the sum of the internal distances of the k nearest neighbours of a_i , denoted as $\hat{B}_i = \sum_{i=1}^n d_{ij}$.
- (7) Calculate the mean value of the distance from a_i to the k nearest neighbours $\bar{B}_{a_i} = B_{a_i} / (\sqrt{k \cdot \hat{B}_{a_i}})$.
- (8) Repeat the above process to iterate through each energy consumption data point to obtain \bar{B}_{a_i} for each energy consumption data, and determine whether B_i is greater than the set distance threshold; if it is greater, the corresponding energy consumption data is considered to be an anomaly; otherwise, it is a normal value, which is noted as y .

After the above steps, the processing of the historical energy consumption data of the building is completed and the values of the output variables of the prediction model are obtained.

4.2. PCA downscaling of factors influencing building energy consumption. In addition to the output variables mentioned above, input variables are indispensable for the establishment of the model, so the acquisition of input variables is crucial. To accurately forecast the energy consumption of buildings, this article performs PCA to reduce the dimensionality of the influencing factors of the building energy system analyzed in Section 3, the principle of which is implied in Figure 2.

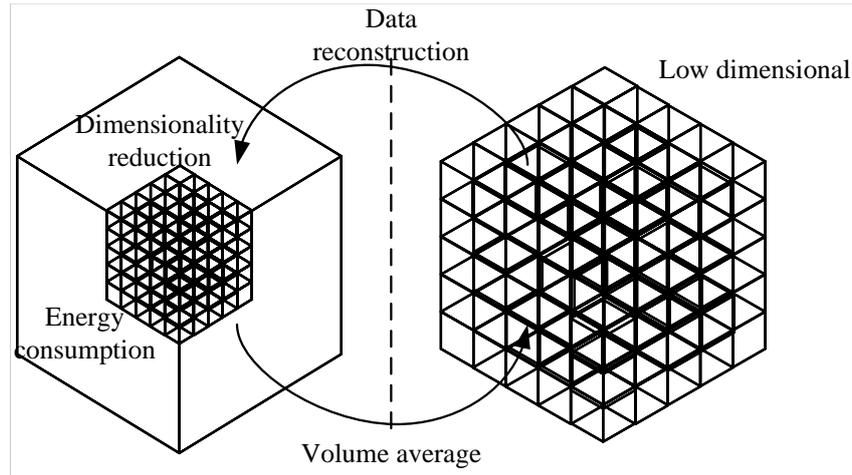


Figure 2. Schematic diagram of PCA downscaling model for energy consumption influencing factors

It is assumed that there are m samples and n characteristics in the factors affecting building energy consumption, which can be expressed as bellow.

$$\begin{cases} x^{(1)}, x^{(2)}, \dots, x^{(m)} \\ x^{(1)} = [x_1^{(1)}, x_2^{(1)}, \dots, x_n^{(1)}]^T \end{cases} \quad (14)$$

During the whole process, it should be ensured that each feature has the same scale, otherwise the whole downscaling effect will be affected. The dimensionality reduction from 3D to 2D can be understood as finding two vectors x_1 and x_2 , constructing a new feature by these two vectors, and corresponding it to the original sampling points. The projection of the sampling points of the three features onto the two-dimensional plane is subject to errors, which must be measured accurately. The formula for calculating the magnitude of the error is implied below.

$$e = \frac{1}{m} \sum_{n=1}^m ||x_n^{(m)} - x_a^{(m)}||^2 \quad (15)$$

where $x^{(n)}$ denotes the mapped coordinate. If the coordinates are represented in the current sampling space, then the dimensions of the downscaling matrix correspond to the sampling values, and a downscaling matrix is constructed based on this.

$$x = \begin{bmatrix} (x^{(1)})^T \\ \vdots \\ (x^{(m)})^T \end{bmatrix} = \begin{bmatrix} x^{(1)} & x^{(1)} & \dots & x^{(1)} \\ x^{(2)} & x^{(2)} & \dots & x^{(2)} \\ \vdots & \vdots & \vdots & \vdots \\ x^{(m)} & x^{(m)} & \dots & x^{(m)} \end{bmatrix} \quad (16)$$

After obtaining the downscaling results, the model is mapped onto a lower dimensional space using a downscaling matrix. The downscaling formula is implied below.

$$x' = (n \times c)^T \cdot (n \times 1) \quad (17)$$

where c denotes the dimensionality of data reduction, the larger the value of c , the more feature vectors in the low dimensional space, the smaller the error of dimensionality reduction and the preservation of the original feature characteristics.

4.3. DSVM-based building energy consumption prediction in a data-driven model. Based on the preparation of the above two chapters, the DSVM model is used to predict building energy consumption with historical energy consumption data y as output and energy consumption influence factors of dimension reduction as input. Firstly, the fractal idea [25] is used to input the segmentation feature vectors into all sub-SVMs for learning to obtain deep features. Then, the idea of feature selection [26] is used to assign weights to the feature vectors in each segment and reorganize the segments to obtain the optimized SVM classifier. Finally, the prediction results of building energy consumption are obtained through the output level.

Suppose there is a training data set $(x', y), i = 1, 2, \dots, n$, and the first level of DSVM data set has m feature vectors $X' = \{X_1, X_2, \dots, X_m\}$, where X_j and the label set associated with the initial feature vector set is $Z = \{z_1, z_2, \dots, z_m\}$. In the input layer of DSVM, X_j is divided into m feature vector segments, the dimension of the segment is N_m , and the corresponding segment feature vectors and labels are implied as bellow.

$$\left\{ \left[\begin{array}{c} X_{i1} \\ \vdots \\ X_{im} \end{array} \right], z_i \right\}, \dots, \left\{ \left[\begin{array}{c} X_{i((m-1)+1)} \\ X_{i((m-1)+2)} \\ \vdots \\ X_{iNm} \end{array} \right], z_i \right\} \quad (18)$$

Subsequently, the feature vectors of the first layer are segmented and the first feature segment X_{ij} is input into SVM_1 for training, and the deeper features of X'_j and their labels are obtained using Equation (19).

$$\{[\beta_1 z_1 L(s_1, X'_j), \beta_2 z_2 L(s_2, X'_j), \dots, \beta_m z_m L(s_m, X'_j)], z_j\} \quad (19)$$

After obtaining the deep-level features of X_{ij} , feature weights are assigned to each SVM_{1j} . According to the relevant principle of SVM [27], the hyperplane that maximizes the distance from the nearest sample point to that hyperplane can maximize the accuracy and generalization ability of classification, and also has better prediction ability for the classification of unknown data. Therefore, in this paper, the idea of feature selection is used to select the maximum distance between all support vectors and hyperplanes of SVM as the key feature weights.

(1) Calculate the maximum distance between each support vector and the corresponding hyperplane for the SVM_{in} classifier in level i .

$$d_{\max} = \sum_{\max} \frac{|v^T x_i + b| C_i}{\|v\| s_i} \quad (20)$$

where v denotes the normal vector to the hyperplane l , b denotes its displacement, s_i denotes the number of support vectors in SVM_{in} , and C_i denotes the total distance from all support vectors to the hyperplane.

$$\begin{cases} C_i = \sum_{i=1}^{r_j} c_i \\ c_i = \frac{y_i(w_j s_i + b_j)}{\|w_j\|} \end{cases} \quad (21)$$

where c_i denotes the hard interval from the support vector to the hyperplane in SVM_{in} , and w_j and b_j are the penalty factor and kernel function parameter of the hyperplane l in SVM_{in} respectively, then the feature weights μ_{1i} of SVM_{in} are implied as bellow.

$$\mu_{1i} = \frac{1}{r_i} \frac{C_i}{d_{\max}} \quad (22)$$

where r_i denotes the minimum number of support vectors.

(2) Each feature in level i is combined again and each feature vector segment is recombined by the features derived in SVM_{in} to obtain the feature vector of the new sample.

$$Y_i^k = \{\mu_{i11}Y'_{i1}, \mu_{i12}Y'_{i2}, \dots, \mu_{i1n}Y'_{iN}\} \quad (23)$$

(3) Segment the input layer using the fractal thinking method, the number of segments is n . In order to ensure that the feature vector segments in the previous layer will not be reclassified, it is necessary to add constraints on n .

$$\frac{mr}{n}/r = \frac{m}{n} = o, o \in Z \quad (24)$$

The k -level DSVM network is trained according to steps (1) to (3), and the output layer is learnt by training using the segmentation of the feature vectors of the previous level to obtain p SVMs, and final classifiers are obtained by combinin

$$\begin{cases} F(x'_i) = \sum_{j=1}^p \alpha_j f_j (\alpha_j \geq 0, \sum_{j=1}^p \alpha_j = 1) \\ y_i = \text{sign}(F(x'_i)) \end{cases} \quad (25)$$

where y_i is the output building energy consumption prediction result, all SVM_{ij} have corresponding weight value α_j , and the accuracy of building energy consumption prediction is determined by Q_j . The larger α_j is, the higher the accuracy is, and $\alpha_j = \frac{Q_j}{\sum_{i=1}^q Q_i}$ the weak classifiers of the output level.

5. Performance testing and analysis.

5.1. Analysis of predicted results. The dataset used in this paper is from RECS [28], a database of building energy consumption statistics provided by the U.S. Department of Energy for model validation and performance evaluation. The dataset consists of offices, university classrooms, university laboratories, libraries, dormitories, and hospitals, and includes two time-dependent variables (i.e., daily and hourly), five weather variables (i.e., air temperature, dew-point temperature, sea level pressure, wind direction, and wind speed), and one building energy data variable (i.e., historical building energy consumption). There are 8792 raw data samples in each building dataset. The dataset is divided into training, test and validation sets in the ratio of 6:3:1. The average of the predictions from 10 runs of the model is taken as the evaluation result. This experiment was conducted based on Pytorch framework with AMD 2700X CPU and NVIDIA RTX 2080 GPU.

The experimental data range from 1 December 2014 to 29 November 2015, a total of 364 days of energy consumption data, the daily energy distribution box plot shown in Figure 3, the daily energy consumption peak distribution in the daytime (8:00–18:00), while the nighttime non-working hours of energy consumption is lower, the energy consumption change is more gentle, indicating that the building energy consumption and the time variable there is a certain correlation.

Based on the experimental data mentioned above, the PB-SVM model in literature [12], the PC-CNN model in literature [18], the PE-SVM model in literature [21] and the model PCA-DSVM in this paper are compared, and the prediction results are implied in Figure 4. The results of 100 test samples of daily energy consumption data are plotted in Figure 4, in which it can be observed that the predicted value of ECP-DSVM is the closest to the actual value, and the overall fluctuation trend is basically the same as the actual value. PB-SVM has a larger error in the peak part of the large changes in energy consumption. This is because PB-SVM predicts building energy consumption based on the traditional SVM algorithm, the training time is long, and the prediction effect decreases with the

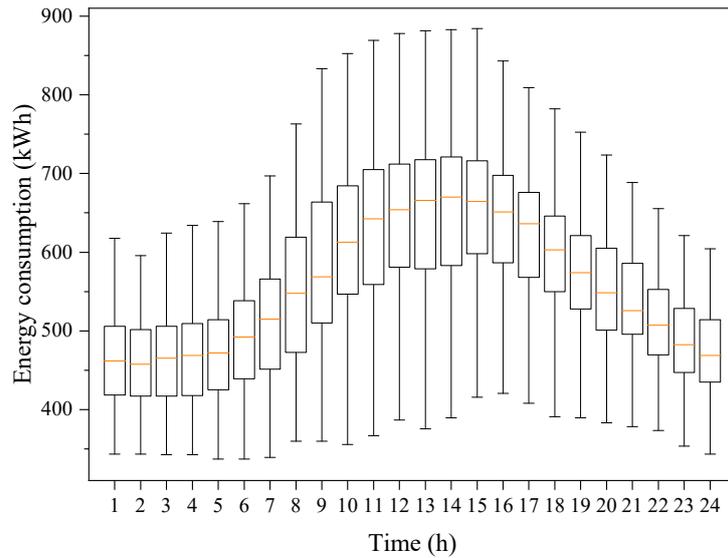


Figure 3. Daily energy distribution box diagram

increase of the sample data volume. PC-CNN has a poorer degree of fitting, due to the fact that the CNN needs to undergo several convolutions before the prediction result can be calculated. PE-SVM uses the linear classifier with DSN architecture for the organisation of SVM to forecast the building energy consumption, and the fitting effect is general. While the ECP-DSVM model has a better fitting degree, relatively good accuracy and robustness, especially in the peak part of the energy consumption that has a large change in energy consumption, which can reflect the law of change of energy consumption more accurately.

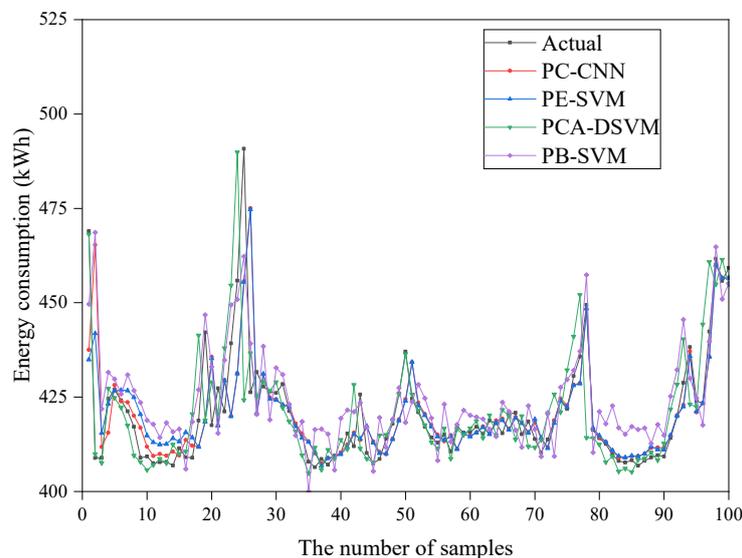


Figure 4. Comparison of building energy consumption prediction results

5.2. Predicted performance comparison. To verify the validity of the suggested model PCA-DSVM, the prediction effect of the comparison models is comprehensively evaluated from five aspects: MAE, RMSE, SMAPE, fitting coefficient R^2 , and maximum relative error E_{max} . Table 1 gives the comparison of the prediction indexes of different

models. As can be seen from Table 1, the designed model ECP-DSVM exceeds other popular prediction methods in all five metrics, where the MAE, RMSE, and SMAPE of ECP-DSVM are 13.48 kWh, 32.59 kWh, and 8.96%, respectively, which are 82.86%, 74.18%, and 79.83% lower than that of PB-SVM, PC-CNN by 74.55%, 62.59%, and 72.53%, respectively, and by 28.25%, 14.01%, and 30.63% than PE-SVM. Furthermore, the R^2 of ECP-DSVM is 0.91, which is 28.17%, 15.19% and 5.81% higher compared with PB-SVM, PC-CNN and PE-SVM respectively, which indicates that the predicted and actual values of ECP-DSVM have a better fitting effect. Comparing the maximum relative error E_{\max} of the four prediction models, it can be seen that the performance of ECP-DSVM is more excellent, which indicates that the selection of air-conditioning energy consumption influencing factors and PCA downscaling have a greater impact on its performance. In summary, ECP-DSVM has more reliable and stable prediction effect, and can provide accurate and effective prediction results for building energy consumption.

Table 1. Comparison of predictive indicators of different models

Model	MAE	RMSE	SMAPE	R^2	E_{\max}
PB-SVM	78.63	126.23	44.43	0.71	47.68
PC-CNN	52.97	87.11	32.62	0.79	33.27
PE-SVM	27.15	54.18	17.24	0.86	20.53
ECP-DSVM	19.48	46.59	11.96	0.91	12.92

Figure 5 shows the time cost of the four-building energy consumption prediction models, PB-SVM takes the longest time to predict building energy consumption, with an average time cost of 150.84 seconds, which is about 19 times that of PCA-DSVM. PC-SVM takes the longest time to predict the energy consumption of a building with an average cost of 150.84 seconds, which is about 19 times that of PCA-DSVM, followed by PC-CNN with an average cost of 92.29 seconds, and PE-SVM with an average cost of 36.51 seconds, which is much shorter than the previous two models, especially Fine-tune. The PCA-DSVM model has the shortest prediction time of 7.95 seconds among the four models.

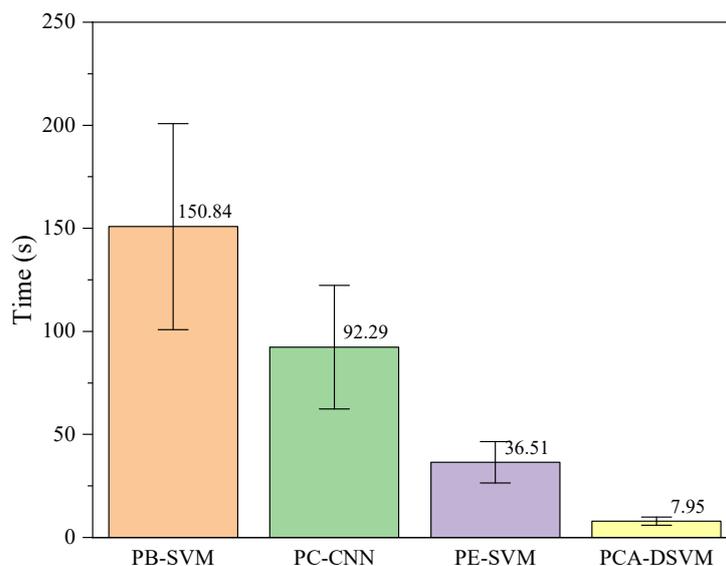


Figure 5. The time cost of the four models

6. Conclusion. As the key target of energy saving and emission reduction strategy, accurate prediction of their energy consumption has a certain guiding significance for the sustainable development of energy. To improve the efficiency of building energy consumption prediction, this article designs a DSVM-based building energy consumption prediction method in data-driven mode. Firstly, a multivariate linear regression model is used to analyze the intrinsic and extrinsic factors of the building energy system, so as to obtain the influencing factors of energy consumption. Secondly, the historical building energy consumption data are preprocessed to eliminate abnormal data, and then the PCA method is used to perform feature dimensionality reduction of the influencing factors to remove redundant features, and then the dimensionality reduced feature vectors are divided into multiple small segments, and the segmented feature vectors are inputted into all the sub-SVMs, to obtain the deep features, and then the feature vectors in the segments are reorganized to be segmented, and the optimized SVM classifiers are obtained. Finally, obtaining the prediction results of building energy consumption through the output level. The experimental outcome implies that the suggested method can accurately predict the building energy consumption with good prediction performance.

REFERENCES

- [1] C. Lu, S. Li, and Z. Lu, "Building energy prediction using artificial neural networks: A literature survey," *Energy and Buildings*, vol. 262, 111718, 2022.
- [2] M. Bourdeau, X. qiang Zhai, E. Nefzaoui, X. Guo, and P. Chatellier, "Modeling and forecasting building energy consumption: A review of data-driven techniques," *Sustainable Cities and Society*, vol. 48, 101533, 2019.
- [3] F. Zhang, T.-Y. Wu, J.-S. Pan, G. Ding, and Z. Li, "Human motion recognition based on SVM in VR art media interaction environment," *Human-centric Computing and Information Sciences*, vol. 9, no. 1, 2019.
- [4] Y. Ma, Y. Peng, and T.-Y. Wu, "Transfer learning model for false positive reduction in lymph node detection via sparse coding and deep learning," *Journal of Intelligent & Fuzzy Systems*, vol. 43, no. 2, pp. 2121-2133, 2022.
- [5] T.-Y. Wu, H. Li, S. Kumari, and C.-M. Chen, "A Spectral Convolutional Neural Network Model Based on Adaptive Fick's Law for Hyperspectral Image Classification," *Computers, Materials & Continua*, vol. 79, no. 1, pp. 19-46, 2024.
- [6] K. Amasyali, and N. M. El-Gohary, "A review of data-driven building energy consumption prediction studies," *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 1192-1205, 2018.
- [7] A. Jahanshahi, D. Jahanianfard, A. Mostafaie, and M. Kamali, "An Auto Regressive Integrated Moving Average (ARIMA) Model for prediction of energy consumption by household sector in Euro area," *AIMS Energy*, vol. 7, no. 2, pp. 151-164, 2019.
- [8] B. Nepal, M. Yamaha, A. Yokoe, and T. Yamaji, "Electricity load forecasting using clustering and ARIMA model for energy management in buildings," *Japan Architectural Review*, vol. 3, no. 1, pp. 62-76, 2020.
- [9] N. Fumo, and M. R. Biswas, "Regression analysis for prediction of residential energy consumption," *Renewable and Sustainable Energy Reviews*, vol. 47, pp. 332-343, 2015.
- [10] K. P. Amber, M. W. Aslam, A. Mahmood, A. Kousar, M. Y. Younis, B. Akbar, G. Q. Chaudhary, and S. K. Hussain, "Energy consumption forecasting for university sector buildings," *Energies*, vol. 10, no. 10, 1579, 2017.
- [11] M. Mohandes, "Support vector machines for short-term electrical load forecasting," *International Journal of Energy Research*, vol. 26, no. 4, pp. 335-345, 2002.
- [12] B. Dong, C. Cao, and S. E. Lee, "Applying support vector machines to predict building energy consumption in tropical region," *Energy and Buildings*, vol. 37, no. 5, pp. 545-553, 2005.
- [13] J. Massana, C. Pous, L. Burgas, J. Melendez, and J. Colomer, "Short-term load forecasting in a non-residential building contrasting models and attributes," *Energy and Buildings*, vol. 92, pp. 322-330, 2015.
- [14] B. B. Ekici, and U. T. Aksoy, "Prediction of building energy consumption by using artificial neural networks," *Advances in Engineering Software*, vol. 40, no. 5, pp. 356-362, 2009.

- [15] W. Zhang, L. Xu, Z. Shen, and B. Ma, "A new approach for mechanical parameter inversion analysis of roller compacted concrete dams using modified PSO and RBFNN," *Cluster Computing*, vol. 25, no. 6, pp. 4633-4652, 2022.
- [16] R. Lei, and J. Yin, "Prediction method of energy consumption for high building based on LMBP neural network," *Energy Reports*, vol. 8, pp. 1236-1248, 2022.
- [17] K. Li, and H. Su, "Forecasting building energy consumption with hybrid genetic algorithm-hierarchical adaptive network-based fuzzy inference system," *Energy and buildings*, vol. 42, no. 11, pp. 2070-2076, 2010.
- [18] Y. Zhao, W. Li, J. Zhang, C. Jiang, and S. Chen, "Real-time energy consumption prediction method for air-conditioning system based on long short-term memory neural network," *Energy and Buildings*, vol. 298, 113527, 2023.
- [19] N. Khan, I. U. Haq, S. U. Khan, S. Rho, M. Y. Lee, and S. W. Baik, "DB-Net: A novel dilated CNN based multi-step forecasting model for power consumption in integrated local energy systems," *International Journal of Electrical Power & Energy Systems*, vol. 133, 107023, 2021.
- [20] H. Yin, Z. Tang, and C. Yang, "Predicting hourly electricity consumption of chillers in subway stations: A comparison of support vector machine and different artificial neural networks," *Journal of Building Engineering*, vol. 76, 107179, 2023.
- [21] M. Savi, and F. Olivadese, "Short-term energy consumption forecasting at the edge: A federated learning approach," *IEEE Access*, vol. 9, pp. 95949-95969, 2021.
- [22] F. Pallonetto, C. Jin, and E. Mangina, "Forecast electricity demand in commercial building with machine learning models to enable demand response programs," *Energy and AI*, vol. 7, 100121, 2022.
- [23] Z. Qi, B. Wang, Y. Tian, and P. Zhang, "When ensemble learning meets deep learning: a new deep support vector machine for classification," *Knowledge-Based Systems*, vol. 107, pp. 54-60, 2016.
- [24] M. Greenacre, P. J. Groenen, T. Hastie, A. I. d'Enza, A. Markos, and E. Tuzhilina, "Principal component analysis," *Nature Reviews Methods Primers*, vol. 2, no. 1, 100, 2022.
- [25] G. Qiang, T. Zhe, D. Yan, and Z. Neng, "An improved office building cooling load prediction model based on multivariable linear regression," *Energy and Buildings*, vol. 107, pp. 445-455, 2015.
- [26] W.-L. Lee, Y.-C. Chen, Y.-C. Chen, and K.-S. Hsieh, "Unsupervised segmentation of ultrasonic liver images by multiresolution fractal feature vector," *Information Sciences*, vol. 175, no. 3, pp. 177-199, 2005.
- [27] B. Venkatesh, and J. Anuradha, "A review of feature selection and its methods," *Cybernetics and Information Technologies*, vol. 19, no. 1, pp. 3-26, 2019.
- [28] Y. B. Dibike, S. Velickov, D. Solomatine, and M. B. Abbott, "Model induction with support vector machines: introduction and applications," *Journal of Computing in Civil Engineering*, vol. 15, no. 3, pp. 208-216, 2001.
- [29] C. Robinson, B. Dilkina, J. Hubbs, W. Zhang, S. Guhathakurta, M. A. Brown, and R. M. Pendyala, "Machine learning approaches for estimating commercial building energy consumption," *Applied Energy*, vol. 208, pp. 889-904, 2017.