Prediction of Hourly Energy Consumption for Office Building Using Optimized SVR

Ziqing Wei[1,a], Ran Xiao[1,b], Gan Yang[1,c] and Xiaoqiang Zhai[1,d]

[1]*Institute of Refrigeration and Cryogenics, Shanghai Jiao Tong University, Shanghai 200240, China*
a. weiziqing@sjtu.edu.cn, b. xiao1ran3@sjtu.edu.cn, c. yanggan666@sjtu.edu.cn,
d. xqzhai@ sjtu.edu.cn

Keywords: Building energy consumption, machine learning, hourly prediction, particle swarms optimization.

Abstract: Energy consumption in buildings takes a large part in the total energy consumptions in modern society. Energy consumption prediction with high precision is necessary for building energy conservation. Meanwhile, due to the difficulty of data acquisition, the influence of model dimensionality reduction on prediction accuracy needs further studies. In this paper, a particle swarm optimization - supported vector regression (PSO-SVR) algorithm is adopted to construct the prediction model. The energy consumption data of an office building in Shanghai from 2017 to 2018 is used to build and test the prediction model. Four different combinations of selected input variables are created and evaluated by the PSO-SVR algorithm to investigate the effect of different input selections. The results show that PSO-SVR algorithm can improve the model performance. The combination investigation shows that the features of indoor facilities influence the prediction performance greater than out door features and time-related features.

1. Introduction

In 2018, energy consumption in buildings took up about 32% of the global energy demand and 30% of energy-related CO_2 emissions. Energy consumption in buildings is still increasing with the population growth and demand for building comfort. Commercial office buildings are an important part of the urban complex. Office buildings normally contain the central heating, ventilation and air-conditioning (HVAC) system, which makes it easier to adjust the system and reduce energy consumption. To achieve the goal, precise prediction of energy consumption is necessary. Precise energy prediction can not only help with the operation scheduling of HVAC system, but also be useful in the energy system designing process.

Black-box model, which is mainly contributed by machine learning algorithm, is widely applied in the prediction of building energy consumption in the past few years. Black-box model ignores the physical relationships in building parameters and gives highly credible value based on statistical theory [1]. Many approaches have been applied in building energy forecasting [2, 3, 4]. Among those approaches, artificial neural networks (ANN) and support vector regression (SVR) got the widespread adoption for their high accuracy [4]. Some further research has focused on the feature selection and tend to reduce the input variables requirement[5, 6, 7], Platon et al [7] used ANN and principal component analysis (PCA) to predict hourly electricity consumption in a government facility. Chen et al [8] adopted SVR algorithm with the multi-resolution wavelet decomposition (MWD) in the short-term electricity consumption prediction.

Meanwhile, some other researches paid attention to the accuracy improvement of prediction model. Almost all of the machine learning algorithms need to setup appropriate hyper-parameters, such as the C, gamma and epsilon in SVR, to insure the fitting (or learning) is effective [9]. However, these parameters have no specific physical meaning and therefore are commonly set with empirical. This is not good enough for improving the precision of the model. Heuristic optimization algorithm can help to set up appropriate parameters, because their unique random search method of optimum can solve the non-convex programming problem within the loss function in machine learning algorithm [10, 11]. Among them, Particle swarm optimization (PSO) shows the strong capability of global optimal search and fast search speed and has been widely used in the hyper-parameter adjustment of machine learning.

Despite the efforts made by the researches mentioned above, difficulties in data collection still cannot be ignored. It is worth to make further investigation about the effects of features reduction on the prediction accuracy.

Considering the research gaps mentioned above, this paper implemented a hybrid PSO-SVR prediction model and used a set of real energy consumption data of a building to validate the improvement. Then the effect of different input variables combination is evaluated by the proposed model.

2. Methodology

2.1.Supported Vector Regression

Support vector machine is firstly created by Vapnik and then extended to support vector regression by introducing the ε-insensitive loss function into the approach. The algorithm is based on the construction of the hyper plane that decided by the maximum margin feature sets. SVR algorithm tries to find out a function that relate the training feature sets $\{\boldsymbol{x_1}, \boldsymbol{x_2}, \boldsymbol{x_3}, \dots\}$and target values $\{\boldsymbol{y_1}, \boldsymbol{y_2}, \boldsymbol{y_3}, \dots\}$. When a new set of features comes up with, the SVR gives the predicted target value based on the function set as follows,

$$f(x) = \langle \boldsymbol{w}, \varphi(\boldsymbol{x}) \rangle + b \tag{1}$$

Where $\boldsymbol{w}$ and b decide the hyper plane, $\varphi()$ is the mapping method on input variables set.

In order to include the training data set as much as possible and keep a high generalization performance of the algorithm simultaneously, the $\boldsymbol{w}$ and b need to be set appropriately. By introducing the ε-insensitive loss function and the slack variables, the SVR problem is converted as follow,

$$\begin{gathered}
\min \frac{1}{2}\|\boldsymbol{w}\|^2 + C\sum_{i=1}^{m}\left(\xi_i + \widehat{\xi_i}\right) \\
\text{s.t. } f(x_i) - y_i \le \varepsilon + \xi_i \\
y_i - f(x_i) \le \varepsilon + \widehat{\xi_i} \\
\xi_i \ge 0, \widehat{\xi_i} \ge 0, i = 1,2,\dots,m.
\end{gathered} \tag{2}$$

where C is the regularization constant, which is also the penalty factor of the objective function, ξ_i and $\hat{\xi}_i$ are the slack variables.

By introducing Lagrange multiplier method, the general formula of the solution is obtained as follows,

$$f(x) = \sum_{i=1}^{m} (\hat{\alpha}_i - \alpha_i)\kappa(\boldsymbol{x}, \boldsymbol{x}_i) + b \tag{3}$$

where $\kappa(\boldsymbol{x}, \boldsymbol{x}_i)$ is called as the kernel function, which can map the non-linear problem to a high dimensional space, and obtain the linearly separable plane. The kernel function that used to deal with the non-linear problem is the Radial Basic Function (RBF), which is suitable for the non-linear mapping between building energy consumption and the inputs. The RBF is described as follow,

$$\kappa(\boldsymbol{x}, \boldsymbol{x}_i) = \exp\left(\frac{\|\boldsymbol{x} - \boldsymbol{x}_i\|^2}{2\sigma^2}\right) \tag{4}$$

Where $\frac{1}{2\sigma^2}$ is called as γ, which can adjust the separation degree of the mapping.

Therefore, the C, ε and γ are the parameters of before SVR algorithm is applied, which are also called as the hyper-parameters. By adjusting hyper-parameters, better performance can be achieved.

2.2.Particle Swarms Optimization

The particle swarm optimization (PSO) was proposed by Eberhart and Kennedy [12] with the inspiration of bird flocking and fish schooling. The PSO algorithm is based on stochastic optimization method with a constant population. The position of each particle in PSO represents a solution to the variables. During the iteration calculation, on one hand, the particles can find out a set of positions by moving itself, the value of best position is called *pbest*. On the other hand, the positions of all particles in the population are compared and the value of global best position, named *gbest*, can be found out. The particles change their position and the velocity based on the *pbest* and *gbest*, and the final position can be obtained by ending up iterations or achieving expected tolerance.

2.3.Research Framework

The research framework is shown in Figure 1. The process starts with data cleaning and normalization, the abnormal data is dropped out from the data set. Then, the independent hour is set as the dummy variable. Moreover, the time series features as HVAC_24, Tmpc_24, Relh_24 are also introduced in to the model. Based on the properties of the features, four combination of input features are set. The data is separated into three parts: The hyper-parameters of SVR model is based on the PSO algorithm with MAPE as the fitness function. A validation is adopted in order to avoid the over fitting problem. The combinations of input are shown in Table 1.

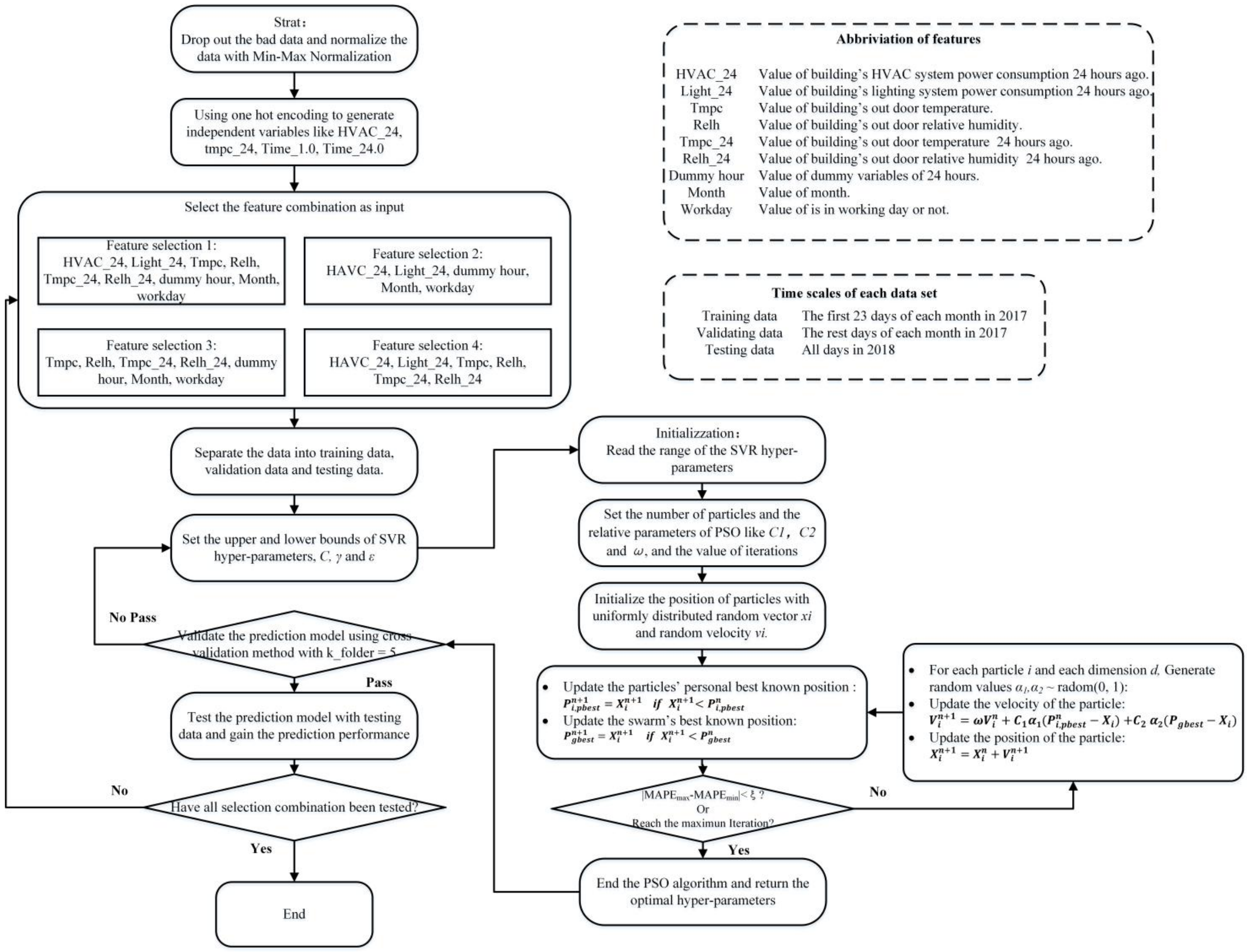

Figure 1: Research framework.

Table 1: Input combinations.

Series	Inputs combinations
Selection 1	HVAC_24, Light_24, Tmpc, Relh, Tmpc_24, Relh_24, dummy hour, Month, workday
Selection 2	HVAC_24, Light_24, dummy hour, Month, workday
Selection 3	Tmpc, Relh, Tmpc_24, Relh_24, dummy hour, Month, workday
Selection 4	HVAC_24, Light_24, Tmpc, Relh, Tmpc_24, Relh_24

3. Case Study

In order to validate the accuracy improvement of PSO-SVR algorithm, the hourly energy consumption data an office building located in Shanghai with corresponding weather conditions are employed in the proposed model. As can be seen from Figure 2, the energy consumption of target building is closely related with the day of the week. The energy consumption shows obvious difference in workday and weekend. The energy consumption of building also shows seasonal changes.

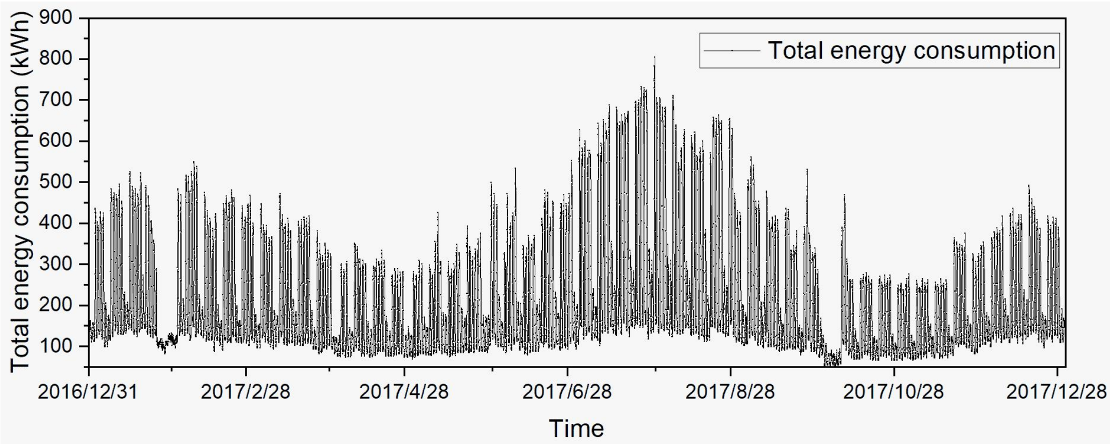

Figure 2: Hourly energy consumption of target building in 2017.

3.1.Data Preparation and Performance Metrics

A filter build by statistics information of energy consumption is adopted to clean the training data. Moreover, linear interpolation method is applied to fulfill the missing data. Then, the max-minimum normalization method is employed to build a stable training data and the testing data by scaling the data to (0 ,1).

To evaluate the prediction model, as shown in Table 2, four performance metrics are used.

Table 2: Detail of performance metrics.

Abbreviation	Description	Formula
MAPE	mean absolute percentage error	$\frac{1}{N}\sum_{i=1}^{N}\frac{\lvert y_i-\hat{y}_i\rvert}{y_i}\times 100\%$
MAE	mean absolute error	$\frac{1}{N}\sum_{i=1}^{N}\lvert y_i-\hat{y}_i\rvert$
RMSE	root mean square error	$\sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i-\hat{y}_i)^2}$
R^2	coefficient of determination	$1-\frac{\sum_{i=1}^{N}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{N}(y_i-\bar{y}_i)^2}$

4. Results

To validate the improvement, a default SVR algorithm with hyper-parameters setting as C=1, γ=0.036, ε=0.1 is implemented. Table 3 shows the performance metrics of original SVR and proposed PSO-SVR algorithm. The MAPE of both algorithms are less than 13% and the R^2 are over 90%, indicating that both algorithms are suitable for building energy consumption prediction. The performance metrics of fitting part are better than the prediction part in both original SVR and PSO-SVR algorithm. Compared with original SVR, proposed PSO-SVR algorithm has the lowest MAPE (8.9323 in fitting part and 9.8032 in prediction part) and the highest R^2 (0.9483 in fitting part and 0.9398 in prediction part), which shows PSO-SVR has improved the performance effectively. The optimal hyper-parameters of SVR are found as C=59.1, γ=0.9421, ε=0.0024.

Table 3: Performance with SVR and PSO-SVR model.

Algorithm	$\{C, \gamma, \varepsilon\}$	Fitting				Prediction			
		MAPE	MAE	RMSE	R^2	MAPE	MAE	RMSE	R^2
SVR	1, 0.036, 0.1	12.5261	0.0332	0.0303	0.9211	11.1396	0.0349	0.0482	0.9129
PSO-SVR	5.91e+02, 9.42e-01, 2.46e-03	8.9323	0.0251	0.0374	0.9483	9.8032	0.0258	0.0276	0.9398

Figure 3 Shows the typical actual and predicted energy consumption of target building. It can be found that the overfitting is more likely to occur in transition season, while in cooling season and heating season, the proposed model shows better generalization ability.

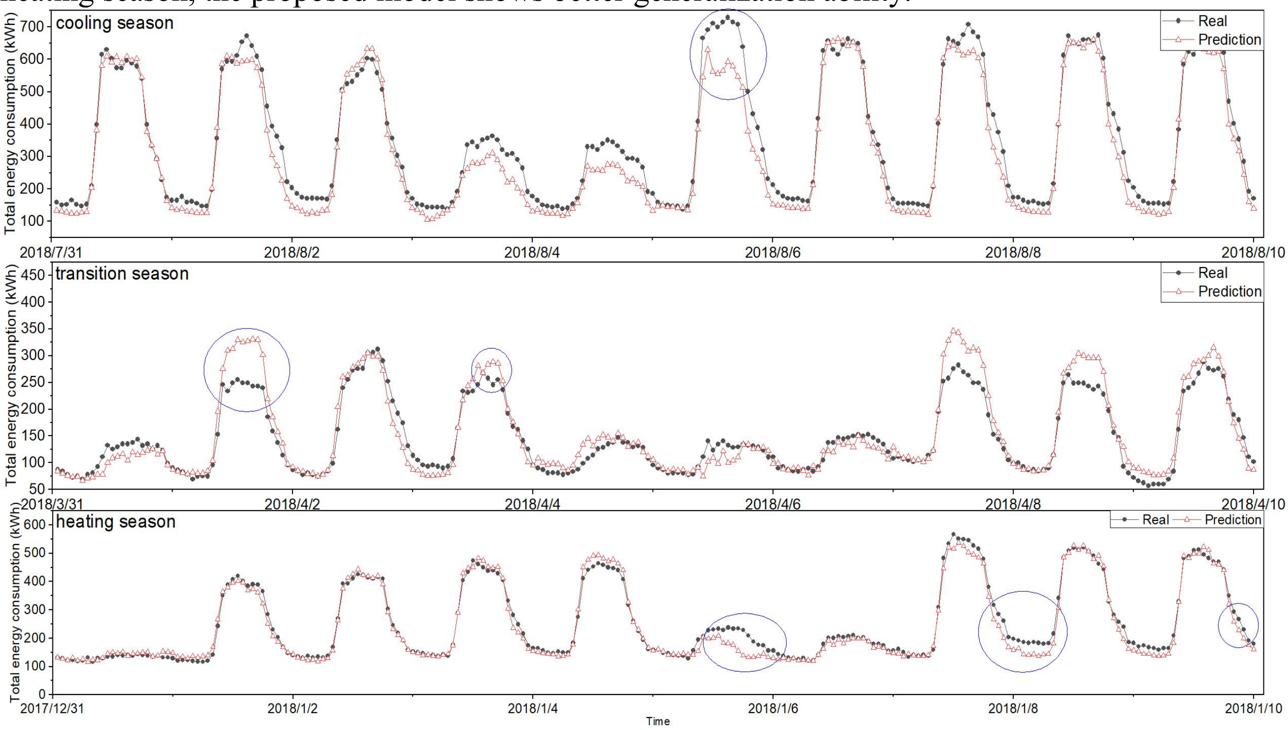

Figure 3: Actual and predicted energy values of target office building.

Table 4: Model performance with different input selection combination.

Algorithm	{C, γ, ε}	Fitting				Prediction			
		MAPE	MAE	RMSE	R^2	MAPE	MAE	RMSE	R^2
Selection 1 (original)	5.91e+02, 9.42e-01, 2.46e-03	8.9323	0.0251	0.0374	0.9483	9.8032	0.0258	0.0276	0.9398
Selection 2	7.71e+02, 6.00e-01, 1.22e-03,	9.2316	0.0298	0.0470	0.9384	10.5514	0.0421	0.0204	0.9201
Selection 3	9.42e+02, 8.36e-01, 2.87e-03,	18.0962	0.0953	0.0887	0.8149	22.7562	0.1044	0.0628	0.7659
Selection 4	8.07e+02, 9.10e-01, 1.12e-03,	9.3625	0.0658	0.0754	0.9146	10.3516	0.0403	0.0278	0.9321

In order to found out the effect of performance degradation on different combination of input features, four combinations are investigated by proposed PSO-SVR model. The optimal hyper-parameters of SVR and corresponding performance metrics of each combination are shown in Table 4. It can be found out that Selection 2 has the lowest performance degradation with MAPE=9.2316 and R^2=0.9384 in fitting part, while Selection 4 has the lowest performance degradation with MAPE=10.3516 and R^2=0.9321 in prediction part. However, performance metrics of Selection 2 decrease obviously, with the obvious overfitting problem in prediction part. This indicates that, when the data of indoor facilities are available, the influence of outdoor environment conditions and features related with time series can only affect the prediction lightly. Therefore, it is necessary to access the data of indoor facilities for high accuracy energy consumption prediction of office building.

5. Conclusions

This paper established a PSO-SVR prediction model for office building's energy consumption. Proposed model has high performance metrics with MAPE=9.8032, MAE=0.0258, RMSE=0.0276and R^2=0.9398. Compared with default model, the MAPE of PSO-SVR prediction model has decreased by 1.33% and the R^2 has increased by 2.69%. During prediction, overfitting problem are more likely occurring in transition season. The influence of indoor facilities features on the prediction results is much greater than outdoor environment features and time-related features. Reducing the indoor facility features as inputs will greatly reduce the accuracy of the model.

Acknowledgments

The study has been supported by the National Key R&D Program of China (Grant No. 2017YFC0704200).

References

[1] Bourdeau M, Zhai X qiang, Nefzaoui E, Guo X, Chatellier P, Modeling and forecasting building energy consumption: A review of data-driven techniques. Sustain Cities Soc. 48(2019) 1015-1033.

[2] Yildiz B, Bilbao JI, Sproul AB. A review and analysis of regression and machine learning models on commercial building electricity load forecasting. Renew Sustain Energy Rev. 73(2017) 1104–1122.

[3] Zhao HX, Magoulès F. A review on the prediction of building energy consumption. Renew Sustain Energy Rev. 16(2012) 3586-3592.
[4] Amasyali K, El-Gohary NM. A review of data-driven building energy consumption prediction studies. Renew Sustain Energy Rev. 81(2018) 1192-1205.
[5] Ding Y, Zhang Q, Yuan T, Yang F. Effect of input variables on cooling load prediction accuracy of an office building. Appl Therm Eng 128(2018) 225-234.
[6] Ding Y, Zhang Q, Yuan T, Yang K. Model input selection for building heating load prediction: A case study for an office building in Tianjin. Energy Build 159(2018) 254-270.
[7] Platon R, Dehkordi VR, Martel J. Hourly prediction of a building's electricity consumption using case-based reasoning, artificial neural networks and principal component analysis. Energy Build 92(2015) 10-18.
[8] Chen Y, Tan H. Short-term prediction of electric demand in building sector via hybrid support vector regression. Appl Energy 204(2017) 1363–1374.
[9] Selakov A, Cvijetinović D, Milović L, Mellon S, Bekut D. Hybrid PSO-SVM method for short-term load forecasting during periods with significant temperature variations in city of Burbank. Appl Soft Comput J 16(2014)800–808.
[10] Gu J, Wang J, Qi C, Min C, Sundén B. Medium-term heat load prediction for an existing residential building based on a wireless on-off control system. Energy 152(2018) 709–718.
[11] Li Y, Wen Z, Cao Y, Tan Y, Sidorov D, Panasetsky D. A combined forecasting approach with model self-adjustment for renewable generations and energy loads in smart community. Energy 129(2017) 216–227.
[12] Eberhart R, James Kennedy. A New Optimizer Using Particle Swarm Theory. Sixth Int Symp Micro Mach Hum Sci 267(1999) 39–43.