STUDY OF THE WHOLE BUILDING ENERGY USE INVERSE MODELING PERFORMANCE THROUGH SUPPORT VECTOR MACHINE REGRESSION

Shinwoo Lee and Juan Carlos Baltazar
Texas A&M University, College Station, TX, USA

ABSTRACT

The performance of a single-variate support vector machine (SVM) was investigated as a whole-building energy use nonlinear inverse modeling tool. Although the SVM is generally employed with multiple attributes, given the benefits of using a single independent variable and for a fair comparison with another conventional building energy inverse modeling method, the change-point regression, only a single attribute was used as an independent variable. Numerical experiments were conducted based on 32 samples of actual chilled water (CHW) and heating hot water (HHW) use in buildings. The outdoor air temperature and outdoor air enthalpy were used as the main regressors. For daily data, although the average performance of SVM models was only slightly better than that of change-point (CP) models, the difference was more remarkable in some samples than in others. However, for monthly data, there was no improvement of performance.

INTRODUCTION

An inverse model is established by making a mathematical or statistical connection between the known dependent and independent variables. The inverse models of building energy systems are used to forecast future energy use, measure and verify the energy savings from retrofitted buildings, and detect and diagnose the faults of building energy systems and their components. For monthly and daily data analysis, steady-state inverse models are useful because transient effects, such as thermal storage, occupancy, and equipment schedules, are eliminated when the data frequency is larger than a daily basis (Kissok et al. 1998). Simple linear regression, multiple linear regression, and CP regression are perhaps the most well-known examples of steady-state inverse modeling, and they are suggested in ASHRAE Guideline 14 and IPMVP as baseline models. Recently, as the use of machine learning or more complicated algorithms have become increasingly user-friendly, many studies have compared the conventional building energy inverse modeling methods with new ones that include different versions of artificial neural networks (ANNs) and SVMs, or have compared other new methods (Ogcu et al. 2012, Rahman et al. 2017, Massana et al. 2015). Some studies have found the performance of SVM to be as high as or higher than that of ANN (Zhao and Magoulès 2012, Li et al. 2009, Zhang et al. 2015).

Although these new and more elaborate methods have generally been determined to produce a better result in terms of accuracy, the conventional method can still be preferred, considering the trade-off between effort and the model's accuracy (Zhang et al. 2015). Moreover, most of the studies that have employed machine learning methods have tended to use multiple independent variables. Understandably, the potential of machine learning methods could be maximized with larger datasets. However, the usefulness of the model may decrease as the number of independent variables increases (ASHRAE Guideline 14). For instance, strong collinearity between independent variables could compromise the reliability of a model. In addition, if any of the independent variable data is not available, the model can neither be established nor used to output the dependent variable. This is perhaps the reason why single-variate models are the most commonly used (ASHRAE 2017).

While a wealth of studies have used simple or multiple linear regression as a reference to demonstrate the performance of the relatively new methods, few studies have used the CP model for this purpose (Zhange et al. 2015, Carpenter et al. 2016). However, in the building energy inverse modeling areas, and particularly if the model is created for the measurement and verification of energy savings, the CP model is as important as the linear regression and has become the best-practice standard. The CP model has advantages on many levels. Firstly, using a single independent variable, the CP model can provide a significantly higher goodness-of-fit than the simple linear regression. Secondly, it provides physical insight into how a building consumes energy, as each parameter of the model has physical meaning. Thirdly, it can cover nonlinear relationships between input and output, to some degree, with its combination of linear pieces. However, despite these advantages, the robustness of the CP model can decrease when the relationship between the independent and dependent variables is increasingly nonlinear. Unfortunately, there are certain elements can intensify this nonlinearity. For example, the efficiency of air conditioning equipment varies for different part-load conditions. Hence, if the change in efficiency is large, the weight of nonlinearity increases (AHSRAE 2017). This nonlinearity can be caused by HVAC system types or settings. For instance, a varying cooling coil set point in a constant volume reheat system can produce a sudden increase in the reheat coil load, causing the heating consumption to be nonlinear (Fu et al., 2020).

As previously stated, considering both the benefits of a single-variate model and the need to tackle nonlinear data, this study aims to determine the potential of a highly flexible single-variate nonlinear inverse model as a whole building energy use modeler. To achieve this goal, SVM is selected for two reasons. Firstly, it is extremely flexible, with almost no geometrical constraints, and is thus appropriate for describing the nonlinear relationship between dependent and independent variables. Secondly, the SVM is relatively easier to implement than the other most popular machine learning method, ANNs, in the sense that it has fewer parameters to adjust (Wang and Srinivasan 2017), while the performance can be similar to or better than that of ANNs. Two types of building energy are used as dependent variables: chilled water (CHW) and heating hot water (HHW). For the modeling of CHW energy use, outdoor air temperature (OAT) and outdoor air enthalpy (OAE) are used as regressors, respectively. For hot water energy use modeling, OAT is used as a regressor. The inverse model in this study is intended to be a steady-state model. In this regard, monthly and daily data are used to train the model.

BACKGROUND

Change-point (CP) Regression

From a statistical perspective, change-point regression is identical to piecewise linear regression. In general, the whole-building energy consumption against the OAT or OAE takes a piecewise linear form, provided that the energy is related to air conditioning. The CP model takes the form of Equation 1 through Equation 3, depending on the energy type and HVAC system of the building.

$$E = b_0 + b_1 (T - b_2)^+ \text{ or } E = b_0 + b_1 (b_2 - T)^+ \quad \text{(Eq. 1)}$$

$$E = b_0 \pm b_1 (b_2 - T)^+ \mp b_3(T - b_2)^+ \quad \text{(Eq. 2)}$$

$$E = b_0 + b_1 (b_2 - T)^+ + b_3(T - b_4)^+ \quad \text{(Eq. 3)}$$

Equation 1 is a mathematical description of a three-parameter change-point (3P-CP) model, while Equation 2 is a four-parameter change-point (4P-CP) model, and Equation 3 is a five-parameter change-point (5P-CP) model. The three-parameter and five-parameter CP models originated from three- and five-parameter Princeton scorekeeping methods (PRISM), which were based on the concept of variable-base degree-day (Fels 1986). Three-parameter CP models account for buildings that show a distinct energy consumption pattern between a weather-dependent and weather-independent range. The point between these two ranges is termed the change point. The 4P-CP model was developed after the 3P-CP model to account for the pattern in which both sides of the change-point are weather-dependent, yet the difference is only the sensitivity to the weather (Ruch and Claridge 1992). Such patterns tend to appear when HVAC systems carry out simultaneous heating and cooling to tackle different thermal loads among different zones, or when air handling units are operating on a variable frequency drive. 5P-CP describes buildings that consume the same energy source for both heating and cooling.

Support Vector Machine (SVM)

SVM is a pattern-recognizing machine learning algorithm whose main applications are classification and regression. It has been applied to building energy modeling since mid 2000s (Dong et al. 2006). SVM regression starts by assuming a linear form of model, regardless of whether the modeler wants the model to be linear or not.

$$y_i = \boldsymbol{\omega} \cdot \boldsymbol{x_i} + b \quad \text{(Eq.4)}$$

The next step is to set the size of a user-defined parameter ε and to find a model f(x) in such a way that the deviation of each data point from the model is less then ε with as small a magnitude of $\boldsymbol{\omega}$ as possible. The magnitude of $\boldsymbol{\omega}$ is interpreted as the flatness of a model. However, there is a chance that such a model does not exist. To address this issue, slack variables ξ and ξ^* are used, which loosen the rule and allow larger than ε deviation of some data points. (See *Figure 1*)

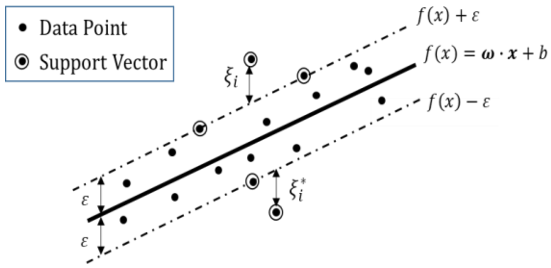

Figure 1 Description of ε-SVM regression procedure (Adapted from Smola and Scholkopf, 2003)

However, data with larger than ε deviation counts as a loss, which is termed ε-insensitive loss (see *Figure 2*), and the regression model is trained in a way that minimizes this loss and the magnitude of $\boldsymbol{\omega}$. This minimization task with a set of conditions takes the form of Equation 5. Mathematically, Equation 5 is equivalent to the convex optimization problem.

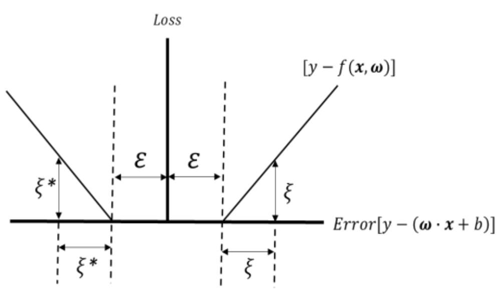

Figure 2 Representation of the ε-insensitive loss function (Adapted from Smola and Scholkopf, 2003)

The strictness of this ε-insensitive loss function can be adjusted by a hyperparameter *C*. For instance, when *C* is small, the minimization function in Equation 5 will add little weight to the ε-insensitive loss term. As a result, the SVM training algorithm becomes insensitive to trivial errors and instead pays more attention to the generalization of the given data. As previously stated, the magnitude of $\boldsymbol{\omega}$ is indicative of the flatness of the model. Thus, Equation 5 shows how SVM handles the trade-off between generalization and accuracy and how it can be adjusted by the user. By using a technique called the kernel trick, the same principle can be applied to train a nonlinear SVM model.

$$Min \ \frac{1}{2}\|\boldsymbol{\omega}\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \qquad \text{(Eq. 5)}$$

$$\text{Subject to} \begin{cases} y_i - (\boldsymbol{\omega} \cdot \boldsymbol{x_i} + b) \leq \varepsilon + \xi_i \\ (\boldsymbol{\omega} \cdot \boldsymbol{x_i} + b) - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases}$$

EXPERIMENT

Input Data

Thirty-two hourly whole-building CHW and HHW uses from university campus buildings from March 2017 to June 2018 were used as input data. One year of data from March 2017 to February 2018 was used as a training dataset and four months of data from March 2018 to June 2018 were used as a testing dataset.

Hourly outdoor dry-bulb temperature and hourly dew point temperature were obtained from the website of the National Oceanic and Atmospheric Administration (NOAA).

For the independent variables, outdoor dry-bulb temperature (OAT) and outdoor air enthalpy (OAE) were used. Three pairs of dependent and independent variables were used for training inverse models: CHW-OAT, CHW-OAE, and HHW-OAT. Since HHW is not affected by latent loads, only OAT was used to create a dataset.

Daily input data and monthly input data were derived from the original hourly data. Monthly data were averaged daily to normalize the impact of different numbers of days per month. Noticeable outliers were manually eliminated to facilitate the analysis of the experiment's results.

SVM Implementation

SVM was implemented by the R programming language with an "e1071" package, which contains the SVM library called LIBSVM (Meyer et al. 2019, Change and Lin 2011). The performance of SVM depends on the selection of a kernel function and hyperparameters. The radial basis function (RBF) was chosen as a kernel, as it generates flexible SVM models. To identify optimal hyperparameters, a one-dimensional grid search was performed sequentially for each hyperparameter. This process was assisted by setting analytically, statistically, or empirically recommended values as initial conditions (see *Figure 3*). Similar logic has been used by previous researchers (Tang et al. 2009, Kaneko and Funatsu 2015). For an initial *C*, a formula proposed by Cherkassky and Ma (2004) was used in the process of statistical reasoning, as described below.

$$C = max\ (\bar{y} + 3\sigma,\ \bar{y} - 3\sigma) \qquad \text{(Eq. 6)}$$

When RBF is used as a kernel function, one more hyperparameter (γ) is added. For an initial γ, an empirically suggested formula was used (Louw and Steel 2006).

$$\gamma = 1/p \qquad \text{(Eq. 7)}$$

Supervised machine learning algorithms generally require a validation process to effectively prevent overfitting. In this study, 10-repeated five-fold cross-validation was used.

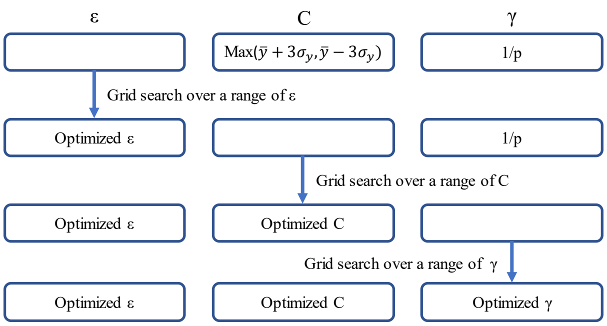

Figure 3 Hyperparameters Selection Procedure

CP Regression Implementation

As a representation of the conventional single-variate regression, a 4P-CP model was used. This is because the HVAC systems of the sample campus building operate continuously throughout the year, for which 4P-CP is generally a better fit than 3P-CP (ASHRAE 2017). 5P-CP was excluded as different energy sources are used for cooling and heating in the sample buildings. To implement 4P-CP, a two-phase grid search algorithm was used (Kissock et al. 2003). In the first phase of the grid search, the whole range of independent variables was tested as a potential CP with a certain resolution. Once one optimal point had been identified, this point was used as the center of the second grid phase. The second phase grid search was performed on a smaller range with a higher resolution. The optimal point identified in the second phase was chosen as the change point.

Evaluation Metrics

Two metrics were used to evaluate the accuracy of the inverse models. Firstly, the coefficient of variance (CV) was used as a measure to assess the deviation of actual data from the SVM or CP models.

$$CV = \frac{\sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(y_i-\hat{y}_i)^2}}{\bar{y}} \times 100 \qquad \text{(Eq. 8)}$$

Secondly, the coefficient of determination (R^2) was used. R^2 is a measure of how much the variance of actual data can be accounted for by the models.

$$R^2 = 1 - \frac{\sum_i(y_i-\hat{y}_i)^2}{\sum_i(y_i-\bar{y})^2} \qquad \text{(Eq. 9)}$$

Paired *t*-test

A *t*-test is a statistical technique to assess whether the difference in means from two different groups is significant or merely the result of chance. A paired *t*-test is a suitable option when the two sample groups are not independent.

DISCUSSION AND RESULT ANALYSIS

SVM Regression Performance

The result of SVM regression on different sets of input data is summarized in Table 1. For CHW, OAE showed approximately 1–2% higher performance as a regressor than OAT.

Table 1 Whole sample mean for different SVM regression models.

Input Attributes	Daily				Monthly			
	Training		Testing		Training		Testing	
	R^2	CV (%)	R^2	CV (%)	R^2	CV (%)	R^2	CV (%)
CHW-OAT	0.937	11.7	0.864	17.5	0.973	7.6	0.969	13.6
CHW-OAE	0.939	11.0	0.872	16.5	0.986	5.2	0.977	12.3
HHW-OAT	0.883	20.7	0.872	26.4	0.973	8.1	0.934	21.8

Table 2 indicates the results of the paired *t*-test between OAT and OAE as a regressor of CHW. The positive mean difference in the table indicates OAE showing a better performance than OAT. According to the paired *t*-test, this difference in performance is not large enough to be considered significant for daily data with 95% confidence. For monthly data, however, the difference in performance can be considered significant, although testing performance in terms of R^2 did not pass the paired *t*-test. The testing performance was worse than the training performance by approximately 5–7%.

Table 2 Paired t-test result between chilled water SVM models based on outdoor air temperature and based on outdoor air enthalpy

Frequency	Data type	Metric	Mean Difference*	t-statistic	Critical t (95% confidence)	Result
Daily	Train	CV (%)	0.70	1.576	2.02	Not significant
		R^2	0.002	0.547	2.02	Not significant
	Test	CV (%)	1.000	1.51	2.02	Not significant
		R^2	0.008	1.298	2.02	Not significant
Monthly	Train	CV (%)	2.4	5.392	2.04	Significant
		R^2	0.013	4.51	2.04	Significant
	Test	CV (%)	1.3	2.175	2.04	Significant
		R^2	0.008	1.521	2.04	Not significant

Comparison with 4P-CP

The results of SVM and 4P-CP for daily data are summarized in Table 3. The average performance of SVM is approximately 0.5% higher than that of 4P-CP.

Table 3 Comparative sample mean performance between SVM and 4P-CP daily regression models

Input Attributes	Training				Testing			
	SVM		4P-CP		SVM		4P-CP	
	R^2	*CV (%)*	R^2	*CV (%)*	R^2	*CV (%)*	R^2	*CV (%)*
CHW-OAT	0.937	11.7	0.931	12.45	0.864	17.5	0.862	17.38
CHW-OAE	0.939	11.0	0.938	11.25	0.872	16.5	0.870	16.82
HHW-OAT	0.883	20.1	0.872	20.7	0.855	26.4	0.852	27.3

Figure 4 and *Figure 5* show a distribution of performance difference between SVM and 4P-CP for CHW-OAE and HHW-OAT, respectively.

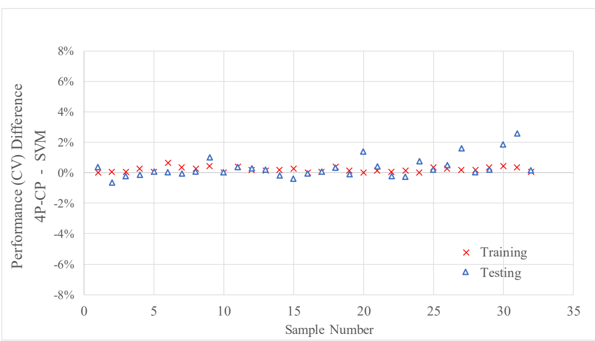

Figure 4. Coefficient of variation difference between 4P-CP and SVM for each sample for daily chilled water use based on outdoor air enthalpy

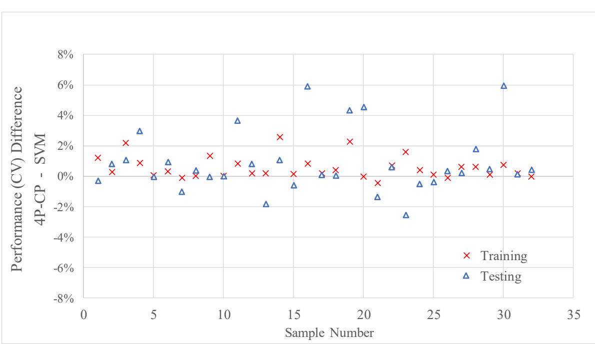

Figure 5 Coefficient of variation difference between 4P-CP and for each sample for daily heating hot water use based on outdoor air temperature

With regard to training performance, SVM showed slightly but consistently higher performance for all types of input attributes. This higher training performance is not markedly different from the expected result given that SVM has an extremely high degree of freedom in fitting while the 4P-CP model is limited by its own geometrical constraints. For testing performance, although 4P-CP performed better in some cases, SVM still produced better results in general. It is noteworthy that the testing performance of some SVM models was significantly higher than that of 4P-CP. Table 4 indicates the results of the paired *t*-test between SVM and 4P-CP regression on daily data. The positive mean difference in the table indicates SVM models performing better than CP models.

Table 4 Paired t-test result between SVM and 4P-CP regression models for daily data

Input Attributes	Data type	Metric	Mean Difference	t-statistic	Critical t (95% confidence)	Result
CHW-OAT	Train	CV (%)	0.74	6.514	2.040	Significant
		R^2	0.007	8.196	2.040	Significant
	Test	CV (%)	-0.15	0.916	2.040	Not significant
		R^2	0.001	0.777	2.040	Not significant
CHW-OAE	Train	CV (%)	0.2	6.911	2.040	Significant
		R^2	0.002	5.115	2.040	Significant
	Test	CV (%)	0.32	2.619	2.040	Significant
		R^2	0.002	2.317	2.040	Significant
HHW-OAT	Train	CV (%)	0.59	4.522	2.040	Significant
		R^2	0.010	4.166	2.040	Significant
	Test	CV (%)	0.89	2.453	2.040	Significant
		R^2	0.003	1.478	2.040	Not Significant

In the training data, the consistently higher performance of SVM was significant according to the paired *t*-test, although the degree of difference in performance was small. In the testing data of the CHW-OAT set, the difference in performance between these two methods was not significant. In the testing data of the CHW-OAE set, the slightly higher performance of SVM was significant according to the paired *t*-test. This higher performance of CHW-OAE in both training and testing datasets implies that OAE could be more appropriate as a regressor for CHW than for OAT, as it reduces the uncertainty caused by latent loads. In the testing data of the HHW-OAT set, the performance of SVM was higher only in terms of CV.

Two samples in which SVM showed significantly higher performance were selected for case analysis. *Figure 6* and *Figure 7* show example plots of such cases.

In *Figure 6*, the original HHW consumption was nearly zero when the OAT was higher than 74 °F. As the outdoor air temperature decreased below 74 °F, the HHW consumption increased in a nonlinear way. SVM successfully captured this mix of linear and nonlinear patterns. Even if the 3P-CP had been used to cover the zero-consumption range, it would not have effectively described the nonlinear range.

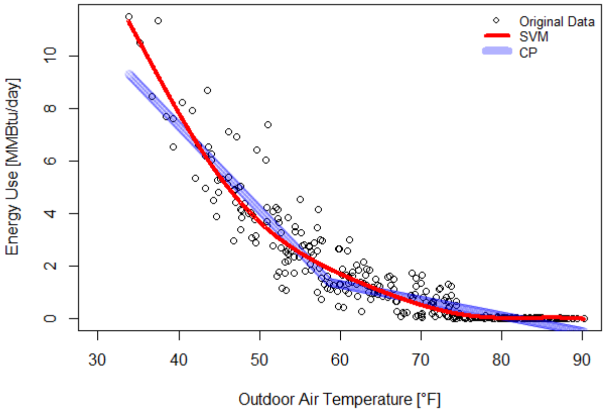

Figure 6 SVM and 4P-CP regression models for daily heating hot water use versus outdoor air temperature for sample 16

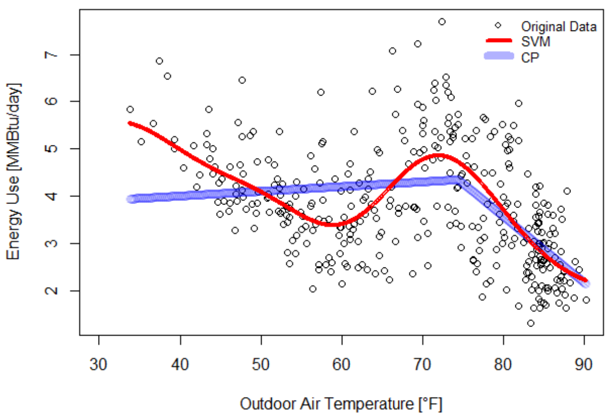

Figure 7 SVM and 4P-CP regression models for daily heating hot water use versus outdoor air temperature for sample 14

In *Figure 7*, the original HHW consumption follows an obviously nonlinear pattern. According to Fu et al. (2020), this pattern appears in constant volume reheat systems with variable cooling coil temperature set points. While the 4P-CP model did not effectively describe the pattern, SVM clearly captured it.

The results of operating SVM and 4P-CP regression on monthly data are summarized in Table 5. Except for the HHW-OAT training dataset, the average performance of 4P-CP was approximately 0.5% higher than that of SVM regression. This result is the opposite of the one from the daily dataset.

Table 5 Comparative sample mean performance between SVM and 4P-CP monthly regression models

Input Attributes	Training				Testing			
	SVM		4P-CPLR		SVM		4P-CPLR	
	R^2	*CV (%)*	R^2	*CV (%)*	R^2	*CV (%)*	R^2	*CV (%)*
CHW-OAT	0.973	7.6	0.977	7.26	0.969	13.56	0.973	12.2
CHW-OAE	0.986	5.21	0.986	5.19	0.977	12.34	0.978	12.06
HHW-OAT	0.958	8.2	0.961	8.7	0.903	21.3	0.936	19.7

Figure 8 shows a distribution of performance for HHW-OAT.

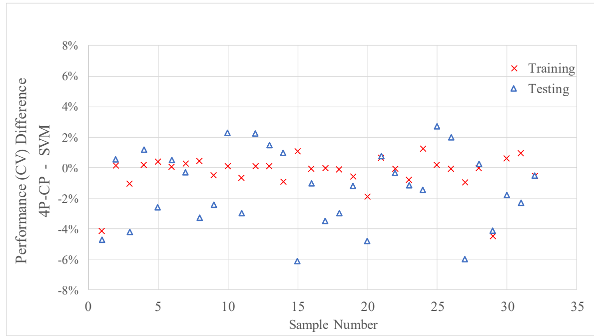

Figure 8 Coefficient of variation difference between 4P-CP and SVM for each sample for monthly average daily chilled water use based on outdoor air temperature.

Unlike its daily data counterparts, SVM did not generally show higher performance in both the training and testing data. The clue about such results can be found in *Figure 9*. Compared to *Figure 6* and *Figure 7*, the regression plots generated by SVM appear to be flatter than those of the 4P-CP models. SVM is a supervised machine learning algorithm; thus, if the amount of training data is insufficient for SVM to make sense of actual patterns, SVM may take a conservative approach, making the regression plot flatter or being far from the expected appearance of physics-based energy models.

Table 6 indicates the results of the paired *t*-test between SVM and 4P-CP regression on monthly data. The positive mean difference in the table indicates SVM models performing better than CP models. In most cases, the difference in performance between the two models was not considered to be significant. There were even three cases in which the performance of SVM was lower than 4P-CPL regression with significance. Thus, the result indicates that SVM is neither practical nor recommended when the given dataset is on a monthly basis.

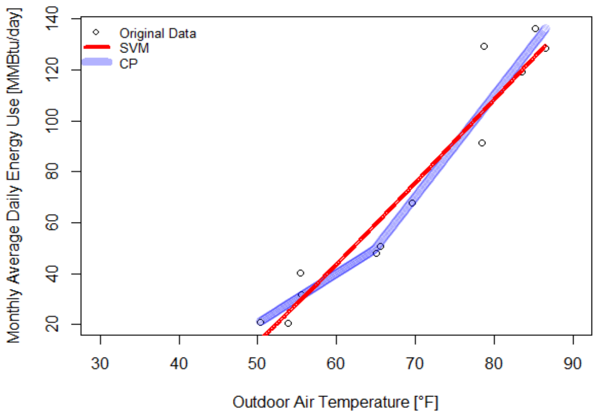

Figure 9. SVM and 4P-CP regression models for monthly average daily chilled water use versus outdoor air temperature for sample 27

Table 6 Paired t-test result between SVM and 4P-CP regression models for monthly data

Input Attributes	Data type	Metric	Mean Difference	t-statistic	Critical t (95% confidence)	Result
CHW-OAT	Train	CV (%)	-0.34	1.568	2.040	Not significant
		R^2	-0.004	3.849	2.040	Significant
	Test	CV (%)	-1.36	3.089	2.040	Significant
		R^2	-0.005	1.094	2.040	Not significant
CHW-OAE	Train	CV (%)	-0.01	0.077	2.040	Not significant
		R^2	-0.001	0.668	2.040	Not significant
	Test	CV (%)	-0.27	0.789	2.040	Not significant
		R^2	-0.001	0.312	2.040	Not significant
HHW-OAT	Train	CV (%)	0.46	1.041	2.040	Not significant
		R^2	-0.003	0.572	2.040	Not significant
	Test	CV (%)	-1.54	2.383	2.040	Significant
		R^2	-0.019	2.346	2.040	Significant

CONCLUSION

The performance of single-variate SVM regression for building energy inverse modeling was investigated.

The whole-building CHW and HHW energy consumption were used as dependent variables. As a regressor for CHW, OAE showed higher performance than OAT, but the degree of difference was not significant enough to pass a paired *t*-test.

The performance of the SVM models was compared with 4P-CP models. When applied to daily data, although the mean accuracy of the SVM models was generally higher than that of the 4P-CP models, this difference was marginal (i.e., less than 1%). To obtain a better understanding of the results, each sample was visually inspected. The visual inspection revealed that even when the CP of the original data was not distinct and the relationship between the dependent and independent variables was closer to a curve than two lines, their numerical accuracy was usually not noticeably affected. This provides an insight into how nonlinearity affects the CP model's accuracy: for a single-variate change-point model, the loss of numerical accuracy due to nonlinearity is generally less than 1%, although the calculated change-point may lose its physical significance. Despite such a marginal mean difference, it is noteworthy that there were some cases in which SVM showed significantly higher performance by as much as 2% and 6% for training and testing data, respectively. This can happen when the weight of nonlinearity is greater than usual.

For monthly data, SVM did not show significantly higher performance than 4P-CP. Rather, the performance of 4P-CP was higher in more cases. This implies that the amount of data was insufficient for SVM to positively make sense of the pattern, resulting in SVM taking a conservative approach.

Overall, the CP model is robust and even generally good at covering nonlinear relationships. Above all, it is also intelligible, as each parameter has physical significance. Thus, when modelers want to regress whole-building cooling or heating energy with a single variable, the CP model would be generally a practical choice, given its accuracy, ease of use, and intelligibility.

However, it would be worthwhile to test SVM or other nonlinear regression methods when the data appears to have a high degree of nonlinearity or when even a small gain in fitting accuracy could be appreciated. In addition, although the running time of SVM is longer than that of the CP regression, the difference would not be large enough to discourage the use of SVM. For instance, to train one daily model of this study, it took less than two minutes with a laptop equipped with Intel (R) Core (TM) i5-8265U CPU @ 1.6Hz and 8GB RAM. As previously stated, the practical benefits of using the SVM are case-dependent. In this study, however, the samples were chosen on a random basis, provided that they had relatively cleaner and more consistent patterns than others. Hence, the focus of future studies should be on clarifying different weights of nonlinearity between different types of building energy systems.

ACKNOWLEDEMENTS

The authors would like to thank Texas A&M University Utility and Energy Service (UES) for their technical and financial support to develop this study.

NOMENCLATURE

E: energy use in buildings

b_0, b_1, b_2, b_3, b_4: parameters of CP-models.

T: outdoor air temperature (°F)

$\boldsymbol{\omega}$: weight vector term of SVM regression

b: bias term of SVM regression

ε: the range of ε-insensitive function where errors are not penalized

C: regularization paramter

γ: hyperprameter of RBF kernel.

$\boldsymbol{x}_i$: indepedent variable for ith observation

ξ_i, ξ_i^*: Slack variable for ith observation

$\bar{y}$: the average of dependent variables

y_i: depedent variable for ith observation

$\hat{y}_i$: modeled value for ith observation

σ: standard deviation

p: the number of independent variables

N: the number of total observations

REFERENCES

ASHRAE Guideline 14, 2014

ASHRAE Handbook - Fundamentals, 2017

Carpenter, J., Woodbury, K., O'Neill, Z., 2016, A Comparison of Gaussian Process Regression and Change-Point Regression For the Baseline Model in Industrial Facilities. ASHRAE and IBPSA-USA SimBuild 2016 Building Performance Modeling Conference. Salt Lake City, UT

Chang, C.,Lin, C., 2011, LIBSVM: A library for support vector machine. ACM Transactions on Intelligent Systems and Technology. 2(3). May 2011 Article No. 27.

Cherkassky, V., Ma, Y., 2004, Practical Selection of SVM Parameters and Noise Estimation for SVM Regression. Neural Networks. 17(1). 113-126.

Dong, B., Cao, C., Lee, S.E., 2006, Applying Support Vector Machines to Predict Building Energy Consumption in Tropical Region. Energy and Buildings. 37(5). 545-553.

Fu, H., Baltazar, J.C., Claridge, D.E., 2020, Break-Point Statistical Model for Building Heating Hot Water Consumption with Constant Volume Reheat Systems. Energy and Buildings. 206. 109575

Kaneko, H., Funatsu, K., 2015, Fast Optimization of Hyperparameters for Support Vector Regression Models with Highly Predictive Ability. Chemometrics and Intelligent Laboratory Systems. 142. 64-69.

Kissock, J.K., Haberl, J.S., Claridge, D.E., 2003, Inverse Modeling Toolkit: Numerical Algorithms. ASHRAE transactions. 109. 425-434.

Kissock, J.K., Reddy, T.A., Claridge, D.E., 1998, Ambient Temperature Regression Analysis for Estimating Retrofit Savings Commercial Buildings. Transactions of ASME. 120(3). 168-176.

Li, Q., Meng, Q., Cai. J., Yoshino, H., Mochida, A., 2009, Predicting Hourly Cooling Load in the Building: A Comparison of Support Vector Machine and Different Artificial Neural Networks. Energy Converstion and Management. 50(1). 90-96.

Louw, N., Steel, S. J.,2006, Variable Selection in Kernel Fisher Discriminant Analysis by Means Of Recursive Feature Elimination. Computational Statistics & Data Analysis. 51(3). 2043-2055.

Massana, J., Pous, C., Burgas, L., Melendez, J., Colomer, J., 2015, Short-term load forecasting in a non-residential building contrasting models and attributes. Energy and Buildings. 92. 322-330.

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F. 2019, e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. R package version, 1.7-2.

Ogcu, G., Demirel, O.F., Zaim, S., 2012, Forecasting Electricity Consumption with Neural Networks and Support Vector Regression. 8th International Strategic Management Conference. Procedia - Social and Behavioral Sciences. 58. 1576-1585.

Rahman, A., Smith, A.D., 2017, Predicting Fuel Consumption for Commercial Buildings with Machine Learning Algorithms. Energy And Buildings. 152. 341-358.

Ruch, D.K., Claridge, D.E., 1992, A 4-Parameter Change-Point Model for Predicting Energy-Consumption in Commercial Buildings. Journal of Solar Energy Engineering. 114(2). 77-83.

Smola, A.J., Schölkopf, B., 2004, A Tutorial on Support Vector Regression. Statistics and Computing. 14. 199-222.

Tang, Y.H., Guo, W.M., Gao, J.H., 2009, Efficient Model Selection for Support Vector Machine with Gaussian Kernel Function. IEEE Symposium on Computational Intelligence and Data Mining.

Wang, Z., Srinivasan, R.S., 2017, A Review of Artificial Intelligence Based Building Energy Use Prediction: Contrasting the Capabilities of Single and Ensemble Prediction Models. Renewable and Sustainable Energy Reviews. 75. 796-808.

Zhang, Y., O'Neil, Z., Dong, B., Augenbroe, G., 2015, Comparisons of Inverse Modeling Approaches for Predicting Building Energy Performance. Building and Environment. 86. 177-190.

Zhao, H., Magoules, F., 2012, A Review on the Prediction of Building Energy Consumption. Renewable & Sustainable Energy Reviews. 16(6). 3586-3592.