Analysis of Conservation Voltage Reduction Effects Based on Multistage SVR and Stochastic Process

Zhaoyu Wang, *Student Member, IEEE*, Miroslav Begovic, *Fellow, IEEE*, and Jianhui Wang, *Senior Member, IEEE*

***Abstract*—This paper aims to develop a novel method to evaluate Conservation Voltage Reduction (CVR) effects. A multistage Support Vector Regression (MSVR)-based model is proposed to estimate the load without voltage reduction during the CVR period. The first stage is to select a set of load profiles that are close to the profile under estimation by a Euclidian distance-based index; the second stage is to train the SVR prediction model using the pre-selected profiles; the third stage is to re-select the estimated profiles to minimize the impacts of estimation errors on CVR factor calculation. Compared with previous efforts to analyze the CVR outcome, this MSVR-based technique does not depend on selections of control groups or assumptions of any linear relationship between the load and its impact factors. In order to deal with the variability of CVR performances, a stochastic framework is proposed to assist utilities in selecting target feeders. The proposed method has been applied to evaluate CVR effects of practical voltage reduction tests and shown to be accurate and effective.**

***Index Terms*—Conservation voltage reduction (CVR), Euclidian distance, short-term load forecasting, Kolmogorov-Smirnov (K-S) test, support vector regression (SVR).**

I. Nomenclature

P_{cvron}	Active load consumption with CVR
P_{cvroff}	Active load consumption without CVR
V_{cvron}	Reduced voltage level
V_{cvroff}	Normal voltage level
P_{pre}	Active load consumption before CVR test
P_{red}	Active load consumption during CVR period
P_{post}	Active load consumption after CVR test
P_{est}	Estimated active load at normal voltage on a CVR day
T_1	Pre-CVR period
T_2	CVR period
T_3	Post-CVR period
ε_{pk}	Euclidian distance-based load index for k th non-test day
P_i	Active load at time i on the test day
P_{ik}	Active load at time i on the k th non-test day
N	Number of data points during T_1 (pre-CVR) and T_3 (post-CVR) periods
ε_{vk}	Euclidian distance-based voltage index for k th non-test day
V_i	Voltage at time i on the test day
V_{ik}	Voltage at time i on the k th non-test day
$\varepsilon_{\mathrm{th}}$	Threshold value of ε_{pk} and ε_{vk}
K	Number of nearest profiles selected by Euclidian distance-based indexes
L_j	Load consumption on day j
$f(\cdot)$	Nonlinear relationship between load consumption and its impact factors
T_j	Temperature profile on day j
H_j	Humidity profile on day j
β	Estimated parameters of MLR
CVR_{fe}	Estimated CVR factor
$\mathrm{CVR}_{\mathrm{fact}}$	Actual CVR factor
A_i	Actual data
F_i	Forecasted data
τ	Forecasting error ($P_e = (1+\tau)P_{\mathrm{act}}$)

Manuscript received May 24, 2013; revised July 09, 2013, August 03, 2013; accepted August 08, 2013. Date of publication September 10, 2013; date of current version December 24, 2013. Paper no. TSG-00416-2013.

Z. Wang and M. Begovic are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: zhaoyuwang@gatech.edu; miroslav@ece.gatech.edu).

J. Wang is with Argonne National Laboratory, Lemont, IL 60439 USA (e-mail: jianhui.wang@anl.gov).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSG.2013.2279836

II. Introduction

CONSERVATION VOLTAGE REDUCTION (CVR) lowers voltages on the distribution system in a controlled manner. CVR can reduce peak demand and achieve more energy savings, while keeping the lowest customer-utilization voltage consistent with levels determined by regulatory agencies and standards setting organizations [1], [2]. CVR is shown to be an established and cost-effective way to reduce peak demand and energy consumption, which has motivated many utilities to investigate its application in individual systems [3]–[6]. How to quantify the effects of CVR has always been a major issue in selecting suitable feeders to implement CVR and performing cost/benefit analysis.

The industry-accepted metric for measuring the performance of CVR is Conservation Voltage Reduction factor (CVRf), which is defined as the percentage of load consumption reduction resulting from one percent reduction in voltage. In this paper, load consumption refers to active power consumed by the load. Previous studies mainly use two methods to calculate the CVR factor, namely the comparison method and the regression method [7], [8]. There are two popular ways to perform the comparison method. The first is to select a group of similar feeders. Reduced voltage is applied to one feeder, while normal voltage is applied to other feeders (control group) at the same time. The second is to perform CVR tests on a feeder and apply normal voltage to the same feeder but during another day with similar weather conditions as the CVR test day. Both comparison methods compare load consumptions of the test and non-test groups to calculate CVR factors. However, it is difficult to find a good 'control group' since there are no two feeders or two days whose operation conditions are exactly the same. The regression method is based on linear regression models that decompose the load, usually into basic and weather dependent components. This method requires long-term "day on/day off" tests and simulates the CVR on/off load for each season based on the assumed model and recorded load as well as weather data. The method can only provide a statistical CVR factor for a certain period such as one season. It is controversial whether this simple linear model can represent complicated load behaviors.

From the engineering point of view, utilities would like to know which feeders are suitable to implement CVR. There are two challenges to answer this question: how to quantify the CVR factor of test feeders; how to select preferred feeders when CVR factors vary from time to time and from feeder to feeder. The major issue in evaluating the CVR effect is to find what the load would be without voltage reduction during the CVR test period. A short-term load forecasting (STLF) method is needed to estimate the normal-voltage load during the CVR period. Many studies have been made in the area of load forecasting and its applications [9]–[16]. Reference [9] applied particle swarm optimization (PSO) and neural networks in load forecasting. Reference [10] used support vector machine (SVM) in mid-term load forecasting (predicting daily load of the next 31 days). Reference [11] combined self-organized map (SOM) and SVM to solve the short-term load forecasting problem (day-ahead prediction). Study in [12] developed a SVR model based on locally weighted vectors for load forecasting. However, these methods are not designed to analyze the CVR effect. Study in [17] used multiple linear regression (MLR) to analyze the impact of demand response (DR) program on energy conservation, load shedding and forecasting at a system level. But the results cannot be used to guide utilities to select suitable feeders to implement DR.

In this paper, a multistage Support Vector Regression-based short-term load forecasting (MSVR-STLF) method is proposed and used to estimate the load consumption at normal voltage levels during the CVR period. As a powerful machine learning method, Support Vector Regression (SVR) is considered as one of the best non-parametric regression techniques, since it can approximate any nonlinear functions [12], [18]. In order to increase the accuracy of the SVR model, only the set of profiles that are close to the load profile under prediction is used to train the SVR model. The selection process is performed by calculating a Euclidian distance-based index in the first stage. SVR is used for load estimation in the second stage. The model accuracy can be largely improved by performing the pre-selection of the training data. To further lower down estimation errors, the estimated profiles are re-selected in the third stage.

A CVR factor is subject to different types of uncertainties, depending on load mix, feeder configurations, weather conditions, human behaviors, etc. For a certain test period, different feeders have different CVR responses. Even for a certain feeder, its CVR factor may vary a lot from one day to another, thus requiring a probabilistic analysis framework. This paper uses the Kolmogorov-Smirnov (K-S) goodness-of-fit test [19] to identify the most suitable probability distributions representing CVR factors of different feeders. The cumulative distribution functions (CDFs) that represent CVR effects of each feeder are used to select candidate feeders. The results could potentially be used to rapidly select target CVR feeders before making any investments.

The main contributions of this paper can be summarized as:

1) Nonparametric load forecasting technique for practical CVR effect analysis at the feeder level based on a large amount of actual field data;
2) A multistage SVR algorithm is developed to improve the forecasting accuracy;
3) A Euclidian distance-based technique is created to validate the results;
4) The probabilistic nature of CVR effects is considered when selecting target feeders to implement voltage reduction.

This paper is organized as follows: Section III introduces CVR and how to evaluate its performance. Section IV details the proposed multistage SVR framework and how to apply it to analyze the CVR effect. The impacts of load estimation errors on evaluation of CVR effects are also presented. The analysis and comparisons of field test results are included in Section V. Section VI presents a stochastic methodology to analyze CVR factors. Section VII concludes the work.

III. Conservation Voltage Reduction

A. Basic Concepts

The CVR effect is evaluated by the Conservation Voltage Reduction factor (CVR_f), which is the relating change in load consumptions to the change in voltage, as defined in (1).

$$\mathrm{CVR}_f = \frac{\%\text{Load Change}}{\%\text{Voltage Reduction}} = \frac{(P_{\mathrm{cvroff}} - P_{\mathrm{cvron}})/P_{\mathrm{cvroff}}}{(V_{\mathrm{cvroff}} - V_{\mathrm{cvron}})/V_{\mathrm{cvroff}}} \quad (1)$$

The major issue in calculating the CVR factor is to estimate P_{cvroff}. Fig. 1 shows a typical CVR test result. The peak demand reduction is achieved by CVR. The test data can be divided into three parts: P_{pre}, P_{red} and P_{post}. The dark bold line represents the measured load. The figure also shows the 'actual' load at normal voltage during the CVR period, which is impossible to know, and the upper and lower bounds of estimated power consumption without CVR.

If the power consumption without CVR could be estimated, the CVR factor can be calculated as follows:

$$\mathrm{CVR}_f = \frac{(P_{\mathrm{est}}(T_2) - P_{\mathrm{cvron}})/P_{\mathrm{est}}(T_2)}{(V_{\mathrm{cvroff}} - V_{\mathrm{cvron}})/V_{\mathrm{cvroff}}} \quad (2)$$

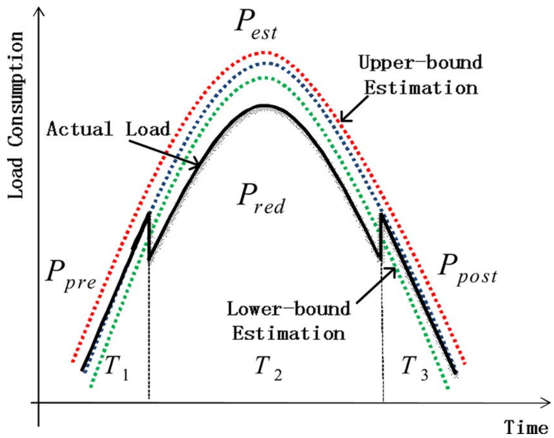


Fig. 1. Typical CVR test result.

In this paper, a multi-stage SVR-based technique is applied to calculate $P_{\rm est}$. $P_{\rm est}$ should fall between the upper and lower bounds, which means there is a certain level of estimation error. The error has impacts on CVR factor calculation, and will be discussed in next section. One important characteristic of the CVR test data is that $P_{\rm pre}$ and $P_{\rm post}$ can be used to find load profiles that are similar to the test profile but at the normal voltage level.

B. Introduction to Field Tests

CVR tests were performed by a utility company on five selected feeders throughout its service area from July, 2011 to March, 2012. Overall, data of 275 days were recorded, of which 120 days were tested. CVR was only implemented during peak hours of these tested days. Measurement devices were installed at each substation to continuously measure kW, kVAR, voltage and current. These data were recorded at one-minute intervals. The devices only recorded data for ten hours out of twenty four-hour operation each day.

IV. Evaluating CVR Effects by Multistage SVR

In order to estimate what the load consumption would be if there is no voltage reduction, the first step is to reconstruct the time series. In this paper, the load is represented by (3).

$$L_j = f(L_{j-1}, L_{j-7}, T_j, H_j) \tag{3}$$

In most cases, a prediction model is trained based on the entire data history, which is called global predictors. However, a better model can be trained by using only the set of points that are close to the point under prediction, which is defined as local predictors [12]. $P_{\rm pre}$ and $P_{\rm post}$ can be used to select load profiles that are similar to the current profile under estimation.

Based on the above analysis, a multistage SVR (MSVR) framework is proposed in this paper and used to estimate the power consumption without CVR. As shown in Fig. 2, measurement data such as power and voltage of both test and non-test days are stored in the database. The rest of the flowchart can be classified into three stages, which represent the following:

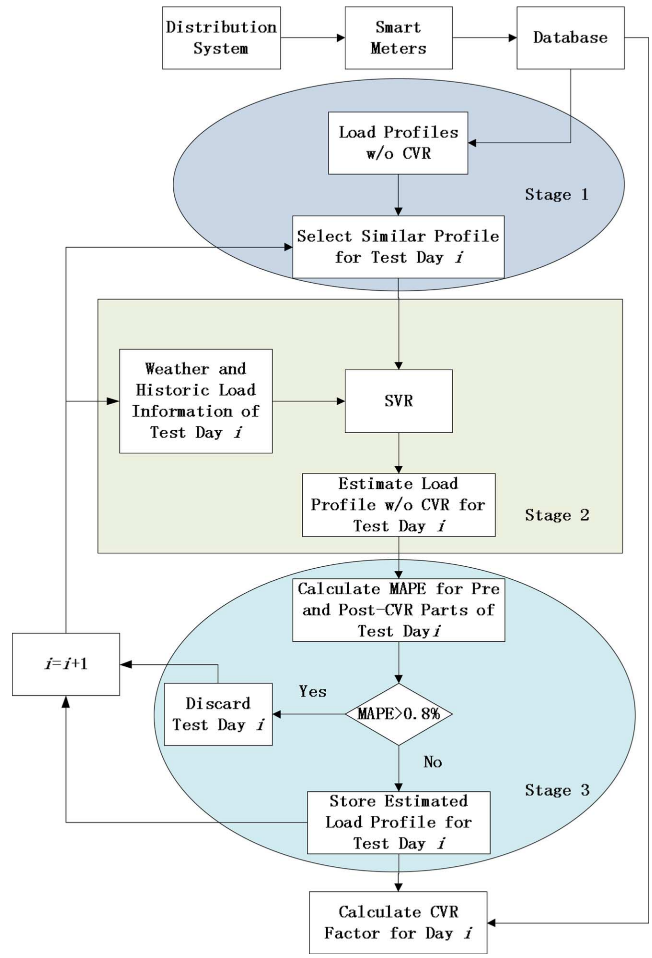


Fig. 2. Multistage SVR framework for CVR analysis.

A. First-Stage: Select Similar Profiles by Euclidian Distance

The purpose of the first stage is to select load profiles from all available historic data to find those are similar to the profile under prediction. This subset of training data is used in local prediction. The similarity between day k and the test day is defined by a Euclidian distance-based index:

$$\varepsilon_{pk} = \sum_{\substack{i=1 \\ i\in T_1,T_3}}^{N} \frac{\sqrt{(P_i - P_{ik})^2}}{\max(P_{ik})\cdot N} \times 100\% \tag{4}$$

Each load profile is divided by the same value which is the largest peak value of all load profiles. By this action, the differences of peak loads of all profiles are maintained. It is clear that the smaller the index, the closer the profile is to the one under prediction. K increases as $\varepsilon_{\rm th}$ becomes larger, which means more dissimilar profiles are included in the training set. In this paper, $\varepsilon_{\rm th}$ is set to be 1%.

B. Second-Stage: Load Consumption Estimation

There are many load consumption estimation methodologies, i.e., multi-linear regression (MLR) [13], [20], [21] and SVR. MLR is one of the most developed and popular forecasting methods. Before comparing the two methods with our dataset, it is necessary to briefly introduce SVR.

Suppose there is a set of training data $\{x_i, y_i\}_{i=1}^n$, where x_i is the input pattern, and y_i denotes the associated output value of x_i. SVR finds a nonlinear map from the input space to the output space and maps the input data to a higher dimensional feature space through this map. Linear regression in the feature space is made by the following estimation function [22],

$$f(x) = \langle \omega, \phi(x) \rangle + b \tag{5}$$

where $\phi(x)$ is the nonlinear mapping from the input space to the high-dimensional feature space, ω denotes the coefficients that need to be estimated, and b is a real constant that also has to be estimated.

The SVR solves an optimization problem [11]:

$$\min \frac{1}{2}\omega^T \omega + C \sum_{i=1}^{n} (\xi_i + \xi_i^*) \tag{6}$$

subject to

$$y_i - \langle \omega, \phi(x_i) \rangle + b \le \varepsilon + \xi_i^* \tag{7}$$
$$\langle \omega, \phi(x_i) \rangle + b - y_i \le \varepsilon + \xi_i \tag{8}$$
$$\xi_i, \xi_i^* \ge 0 \tag{9}$$

where ξ_i^* is the slack variable of the upper training error (ξ_i is the lower one) subject to the ε-insensitive tube $|y - (\langle \omega, \phi(x) \rangle + b)| \le \varepsilon$, The constant $C > 0$ determines the tradeoff between the flatness of f and its accuracy in capturing the training data.

The constraints of (7)–(9) imply that most of the data x_i are placed inside the tube ε. If x_i is outside the tube, there is an error ξ_i or ξ_i^* that needs to be minimized in the objective function. SVR avoids underfitting and overfitting of the training data by minimizing the regularization term $\omega^T \omega$ as well as the training error $C\sum_{i=1}^{n}(\xi_i + \xi_i^*)$.

In the solving process of SVM, Lagrange multipliers α_i and α_i^* are introduced, the SVR training procedure is to solve the dual problem of (6):

$$\min_{\alpha,\alpha^*} \begin{cases} \frac{1}{2}\sum_{i,j=1}^{n} (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) Q(x_i, x_j) \\ +\varepsilon \sum_{i=1}^{n} (\alpha_i + \alpha_i^*) - \sum_{i=1}^{n} y_i (\alpha_i - \alpha_i^*) \end{cases} \tag{10}$$

subject to

$$\sum_{i=1}^{n} (\alpha_i - \alpha_i^*) = 0 \tag{11}$$
$$0 \le \alpha_i, \alpha_i^* \le C \tag{12}$$

where $Q(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle$ is the kernel function. In this paper, the Gaussian kernel as defined in (13) is used.

$$Q(x_i, x_j) = \exp(-\|x_i - x\|^2 / 2\sigma^2) \tag{13}$$

The SVR output is:

$$\hat{f}(x) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) Q(x_i, x) + b \tag{14}$$

In this paper, we define the above algorithm as single SVR. As discussed in Section III, there are 155 non-test days; data of these days are used to evaluate the performance of each method. 55 non-test days are randomly selected to be validation data, data of the remaining 100 non-test days belong to the training set. Load is represented as shown in (3). The training set is used to train the MLR and SVR models. Details about MLR can be found in [13], [21]. The validation set is used to show the performances of the trained models. The MLR model used in this paper is defined as

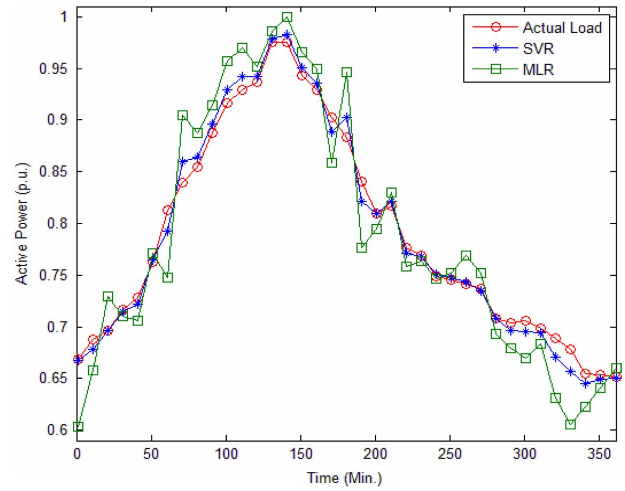


Fig. 3. Actual and forecasted load of feeder 1 on Feb. 14, 2012.

TABLE I
ESTIMATION ERRORS OF SINGLE SVR AND MLR

	MAPE	REP	NMSE
Single SVR	1.20	1.48	0.02
MLR	3.81	4.22	0.06

$$\begin{aligned} \hat{L}_j = \beta_0 &+ \sum\nolimits_{l=1,3,7} \beta_{1,l} L_{j-l} + \sum\nolimits_{l=0,1,7} \beta_{2,l} T_{j-l} \\ &+ \sum\nolimits_{l=0,1,7} \beta_{3,l} T_{j-l}^2 + \sum\nolimits_{l=0,1,7} \beta_{4,l} T_{j-l}^3 \\ &+ \sum\nolimits_{l=0,1,7} \beta_{5,l} H_{j-l} \\ &+ \sum\nolimits_{l=0,1,7} \beta_{6,l} H_{j-l}^2 + \sum\nolimits_{l=0,1,7} \beta_{7,l} H_{j-l}^3 \end{aligned} \tag{15}$$

Fig. 3 shows estimation results of feeder 1 on a day in February, 2012. It can be seen that the SVR model developed in this paper has a better performance than the MLR benchmarking model as specified in (15). The estimation performance is quantified with the mean absolute percentage error (MAPE), normalized mean square error (NMSE) and relative error percentage (REP). The definitions of MAPE and NMSE can be found in [18]. REP is defined as

$$\text{REP} = \sqrt{\frac{\sum_i (A_i - F_i)^2}{\sum_i A_i^2}} \times 100 \tag{16}$$

Estimation errors of the days in the validation dataset are averaged and shown in Table I. It can be seen that the SVR model provides significantly better estimation than the MLR model.

Thus, SVR is used in the second stage for time-series learning and prediction. The K selected profiles are used to train the SVR model.

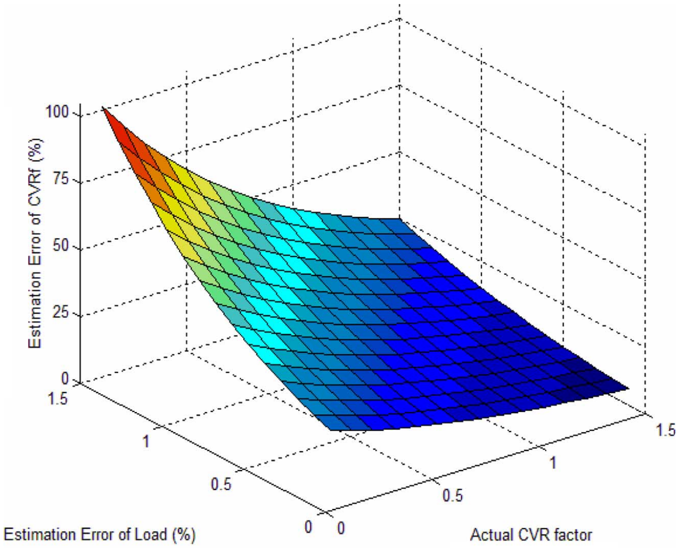


Fig. 4. Relationship of estimation errors of load, CVR factor and estimation errors of CVR.

C. Third-Stage: Re-Select Load Profiles

By taking advantage of the pre-selecting step, the proposed method is expected to have lower errors. However, since the CVR effect is generally a few percentages, the accuracy of CVR effect estimation is highly dependent on the accuracy of estimated load. Equations (17a)–(17d) show how forecasting errors impact the accuracy of CVR effect estimation. In (17d), if there is no forecasting error ($\tau = 0$), CVR_{fe} equals $\mathrm{CVR}_{\mathrm{fact}}$. Though $\mathrm{CVR}_{\mathrm{fact}}$ is unknown, it can be seen that as $|\tau|$ becomes larger, CVR_{fe} differs more from $\mathrm{CVR}_{\mathrm{fact}}$. Moreover, the impacts of $|\tau|$ will be enlarged if the CVR effect is small (γ approximates 1).

$$\mathrm{CVR}_{fe} = \frac{(P_e - P_{\mathrm{cvron}})/P_e}{(V_{\mathrm{cvroff}} - V_{\mathrm{cvron}})/V_{\mathrm{cvroff}}} \tag{17a}$$

$$\mathrm{CVR}_{\mathrm{fact}} = \frac{(P_{\mathrm{act}} - P_{\mathrm{cvron}})/P_{\mathrm{act}}}{(V_{\mathrm{cvroff}} - V_{\mathrm{cvron}})/V_{\mathrm{cvroff}}} \tag{17b}$$

$$P_e = (1+\tau)P_{\mathrm{act}}, \gamma = P_{\mathrm{red}}/P_{\mathrm{act}} \tag{17c}$$

$$\frac{\mathrm{CVR}_{fe}}{\mathrm{CVR}_{\mathrm{fact}}} = 1 + \tau + \frac{\tau}{1-\gamma} + \frac{\tau}{1-\gamma} \tag{17d}$$

Fig. 4 demonstrates the relationship among the MAPEs of estimation, assumed actual CVR factors and errors of CVR effect estimation. It can be seen that lower CVR factors and higher MAPEs will result in larger errors of the estimated CVR factors. In this stage, the estimated load profiles are re-selected to further lower down the estimation errors. As it is clear that P_{cvroff} is unknown on a test day, the MAPEs between $P_{\mathrm{est}}(T_1, T_3)$ and $P_{\mathrm{pre}}, P_{\mathrm{post}}$ are used for re-selection. If the MAPE is smaller than 0.8, P_{est} is stored for further analysis, otherwise, it is discarded.

In order to show the performance of the proposed MSVR model, we calculate the relative errors (REs) between forecasted loads and actual loads in the validation set. The RE of an estimation point i can be defined as

$$\mathrm{RE}_i = \frac{A_i - F_i}{A_i} \tag{18}$$

The mean of REs is 0.134, and the variance is 0.0692.

TABLE II
ESTIMATION ERRORS OF MSVR AND SINGLE SVR

	MAPE	REP	NMSE
MSVR	0.61	0.72	0.01
Single SVR	1.54	1.61	0.03

V. NUMERICAL STUDIES

The proposed multistage SVR method is applied to analyze practical CVR tests. The typical test data for a day can be divided into three parts as discussed in Section III. The devices only recorded ten-hour instead of 24-hour operation data each day, which is sufficient for training the SVR model since CVR tests were only performed during peak hours. CVR factors vary from time to time and from circuit to circuit, so the proposed method is applied to the data of each feeder on all test days. We also use the Euclidian distance-based comparison method to validate our MSVR-based approach.

A. Methodology Verification

It is clear that the data of CVR test days cannot be used to verify whether the proposed method is better than single SVR since the load at normal voltage is unknown. For this reason, one non-test day in each month from July, 2011 to March, 2012 is selected and used to evaluate the performance of the proposed method. Since there are five feeders and nine test months, forty-five datasets in total are used to verify the proposed method. For each dataset, the data covers 360 minutes. Estimation errors of the forty-five datasets are averaged and shown in Table II. It can be seen that the MSVR provides significantly better estimation than single SVR.

B. Estimation of CVR Effects for an Example Feeder

In order to show how the proposed method works to evaluate the CVR effect, the test data of Feeder 1 on a winter day is selected as an example. As shown in Fig. 5, CVR starts at 140 minutes and ends at 420 minutes; this part of data is defined as P_{red} in Section II. The first 140-minute data and the last 180-minute data are defined as pre-CVR (T_1) and post-CVR data (T_3), respectively. They are used in the first stage of the MSVR framework to select sets of non-tested load profiles that are similar to the current profile under prediction. $\varepsilon_{\mathrm{th}}$ is set to be 1% and K is 63. The selected profiles are used to train the SVR model in the second stage. Finally, the load consumption without CVR on that test day is estimated by the trained model. The estimated load and the load data with CVR are shown in Fig. 5.

The Euclidian distance-based comparison method is developed to validate the CVR factor calculated by the proposed method. The method calculates ε_p as defined in (4) and ε_v as defined in (19) so as to find a non-test day whose load and voltage profiles are close to those of the test day.

$$\varepsilon_{vk} = \sum_{\substack{i=1 \\ i \in T_1, T_3}}^{N} \frac{\sqrt{(V_i - V_{ik})^2}}{\max(V_{ik}) \cdot N} \times 100\% \tag{19}$$

CVR factors can be then calculated by comparing the load and voltage differences of the non-test and test days as shown in (2).

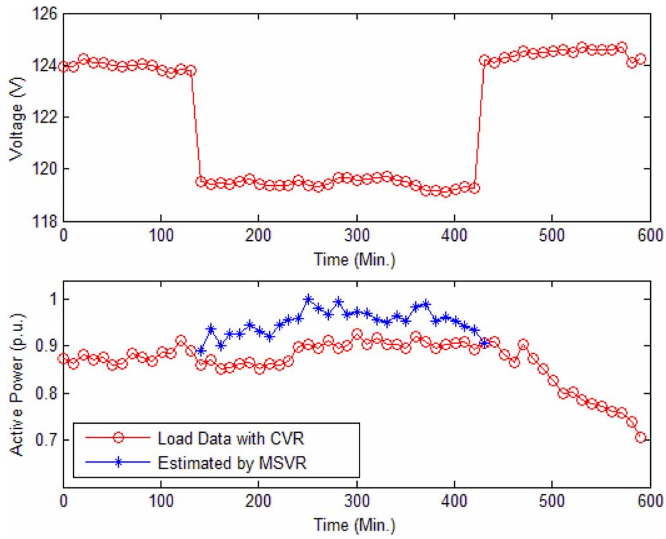


Fig. 5. Voltage profile, actual load profile (w/ CVR) and estimated load profile on a winter day.

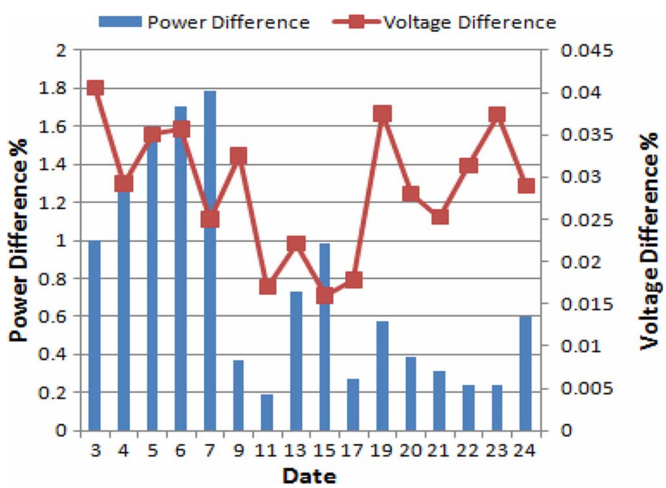


Fig. 6. Calculated Euclidian distance-indexes of Feeder 1 on all non-test days in the same winter month.

Since Feeder 1 on a winter day is used as an example in the previous session, it is necessary to select a 'control group' from all load and voltage profiles of Feeder 1 on non-test days in the same winter month. Control group refers to the load and voltage data under CVR. Fig. 6 shows all the calculated power differences (ε_p) and voltage differences (ε_v). Date 11 is the most similar to the test day and is used as a 'control group' to calculate the CVR factor. Fig. 7 shows CVR factors calculated by the MSVR and the comparison method.

The line with circle markers represents the CVR factors calculated using MSVR estimation. The line with stars represents the CVR factors derived from the control group-based comparison method. It can be seen that the comparison method is similar to the estimated CVR line, which can validate the accuracy of the proposed method. However, as there is no guarantee that a good control group always exists for any feeder on any test day, the proposed MSVR method is more effective. In case the good control group is not available, we cannot validate the calculated CVR factors, which is the limitation of the Euclidian distance-based comparison method. For the 120 test days in this paper, a good 'control group' exists for 46 days. The average MAPE representing the differences between CVR factors calculated by the multistage SVR and the comparison method is 0.69%, which shows the accuracy of the proposed method.

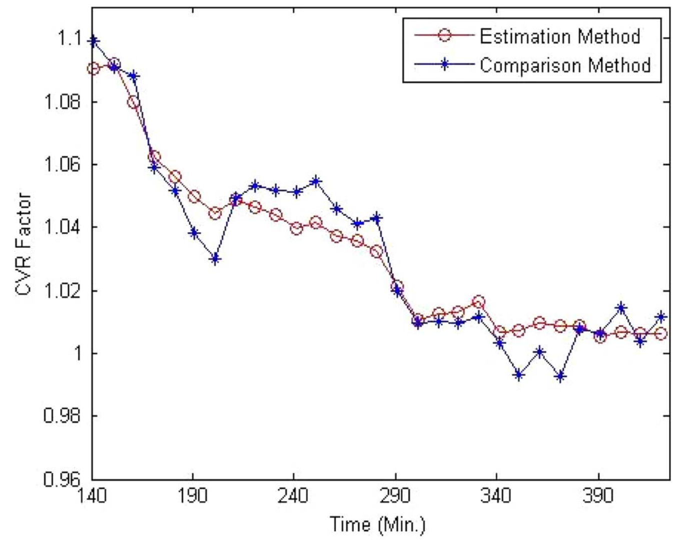


Fig. 7. CVR factors calculated by SVR and the comparison method.

TABLE III
RESULTS OF CVR FACTOR CALCULATIONS OF FIVE FEEDERS

No.	Summer Day Measurement				Winter Day Measurement			
	B	E	M	S/M	B	E	M	S/M
1	0.77	0.89	0.82	8.04	1.00	1.09	1.04	4.03
2	0.90	0.96	0.93	3.21	1.10	1.18	1.15	3.92
3	0.82	0.90	0.85	5.93	1.08	1.15	1.12	3.51
4	0.86	0.95	0.91	5.12	1.21	1.31	1.27	4.33
5	0.68	0.76	0.71	7.43	1.01	1.12	1.07	5.71

C. Overall Results

The overall results of all five feeders on a summer day and a winter day are summarized in Table III, where B represents the beginning value of the CVR factor, E represents the ending value, M denotes the mean value of the CVR factor during the test period and S/M is the standard deviation of the CVR factor divided by its mean value and is represented in percentage.

It can be shown from Table III that CVR factors are not constant but always fluctuating and tend to decrease during test periods. Therefore, continuous monitoring and real-time CVR factor calculation are necessary. For a certain day, CVR effects are different from one feeder to another while a certain circuit's CVR effects are quite different from one day to another. This result indicates that the CVR effects must be evaluated circuit by circuit and day by day. Also, it is clear that CVR factors in winter are much higher than those in summer, which might be due to the large amount of resistive heating loads in winter.

Fig. 8 shows the box plot of CVR factors on all test days from July, 2011 to March, 2012. Boxplot shows the maximum value, minimum value, median, upper quartile (75% quartile), lower quartile (25% quartile) and outliers of the data. As shown in Fig. 8, the maximum value of CVR factor of feeder2 is 1.29,

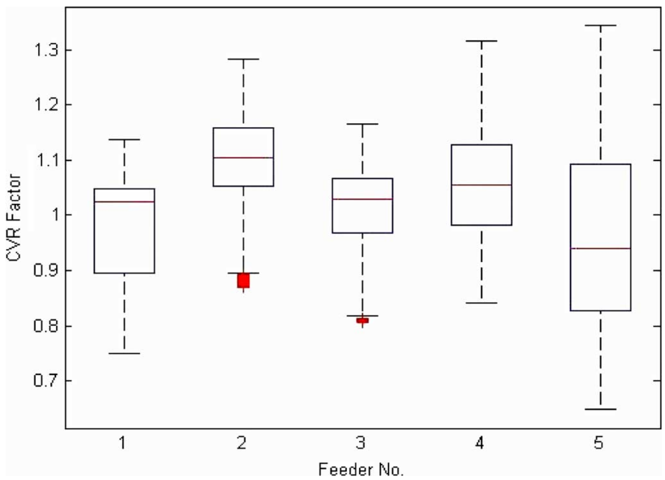


Fig. 8. CVR factors of all feeders on all test days.

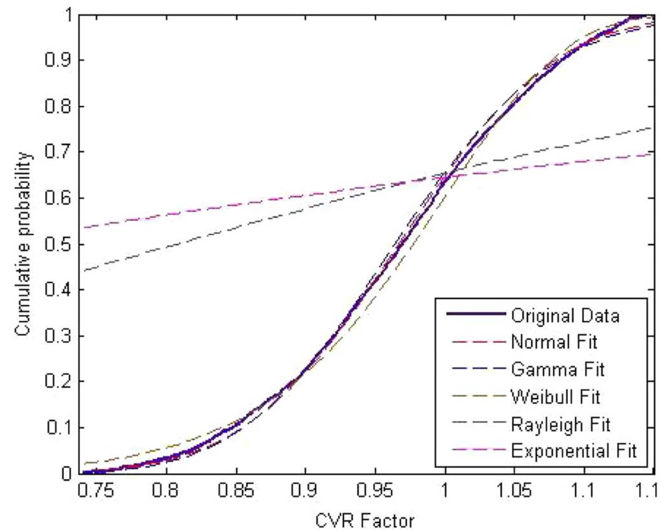


Fig. 9. CDFs of CVR factors of Feeder 1 and various probability distributions.

the minimum value is 0.89, values that are smaller than 0.89 are identified as outliers. Feeder 2 has the largest median CVR factor and Feeder 5 has the smallest one. However, Feeder 5 has the largest CVR factor. This indicates Feeder 5 has a large variance of CVR performances. Since CVR factors may vary a lot from one day to another, preferred CVR feeders cannot be selected by comparing only the statistical median or mean value of CVR factors of each feeder. It can be seen from the boxplot that the fluctuations of CVR factors are relatively large and different from feeder to feeder. This increases the difficulty of selecting target CVR feeders.

VI. Stochastic Analysis

Because of the variability of CVR factors, the CVR effect of each feeder cannot be evaluated deterministically but determined probabilistically. The target CVR feeders can be selected by comparing their probabilistic CVR performances.

In order to identify the most suitable probability distributions for CVR factors of each feeder, the Kolmogorov-Smirnov (K-S) goodness-of-fit test has been carried out [19]. The K-S test computes the test error ψ, which is the maximum vertical distance between a sample cumulative distribution function (CDF) and a fitted CDF. This error is compared to a critical value ψ_{crit}, and any probability distribution fit that satisfies $\psi \leq \psi_{\mathrm{crit}}$ could be accepted.

Fig. 9 shows the differences between the CDF of CVR factors of Feeder 1 and various other CDFs (normal, gamma, Weibull, Rayleigh and Exponential). It is clear that the normal distribution exhibits the most promising goodness-of-fit. Table IV shows the K-S test errors subject to a normal distribution and the maximum likelihood estimates for parameters. μ is the mean and σ is the standard deviation, $\psi_{\mathrm{crit}} = 0.0258$ for the normal distribution fit with a level of significance 5%.

Fig. 10 shows the CDF chart of CVR performances of all feeders. The CDF gives the probability that the variable CVR_f takes a value less than or equal to some specified value CVR_{f-spc}. Table V summarizes CVR effects of all test feeders. It shows the percentiles, which represent the certainty level of achieving a CVR factor below a particular threshold.

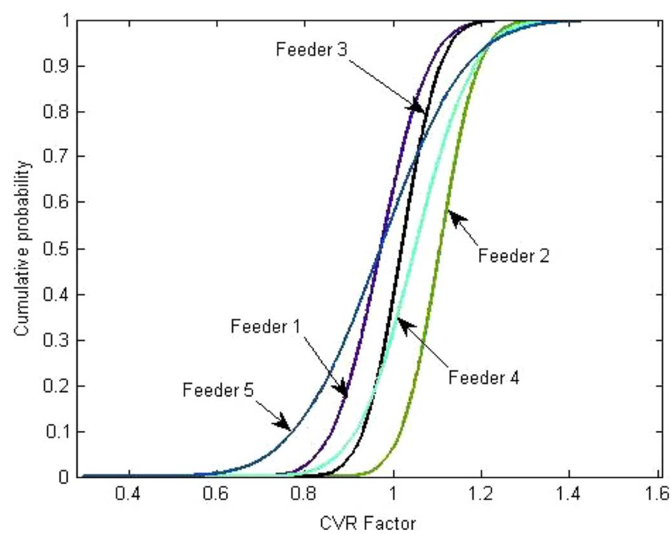


Fig. 10. CDFs of CVR factors of all test feeders.

TABLE IV
K-S Test Errors and Estimated Parameters

Feeder No.	ψ	μ	σ
1	0.0205	0.9716	0.0868
2	0.0122	1.1061	0.0697
3	0.0195	1.0191	0.0687
4	0.0209	1.0503	0.1056
5	0.0185	0.9702	0.1532

TABLE V
Summary of CVR Factors of All Test Feeders

Percentile	0%	25%	50%	75%	100%
$CVR_{f\,\max}$	0.9444	1.0593	1.1062	1.1531	1.3279
$CVR_{f\,\min}$	0.6137	0.8673	0.9708	1.0303	1.1737

The $\mathrm{CVR}_{f\,\max}$ and $\mathrm{CVR}_{f\,\min}$ represent maximum and minimum CVR factors at different percentile levels. For all five test feeders, there is zero chance that any of the feeder exhibits a CVR factor less than 0.6137.

If there are no intersections among the CDF curves, the CDF on the far right of the CDF chart offers the best opportunity for achieving the highest CVR factor at every confidence level, and this feeder is the best CVR candidate. If the CDF curves intersect (as they are in this example), the best feeder is the one that gives the highest CVR factor with the predefined certainty level. If the certainty level is defined to be 90%, then it is clear that Feeder 2 is the best candidate and Feeder 1exhibts the worst performance.

VII. CONCLUSION

In this paper, we proposed a multi-stage SVR technique to estimate the load consumption without voltage reduction during a CVR period. The first stage is to make full use of pre-CVR and post-CVR data to calculate a Euclidian distance-based index, and to select a set of load profiles that are closest to the profile under estimation. The selected profiles are used to train the SVR prediction model in the second stage. Estimated load profiles with large errors are filtered out in the third stage. The CVR factor can be calculated by using the estimated load profile. The impacts of load estimation errors on CVR factor calculation are analyzed. The multistage SVR framework is validated by the Euclidian distance-based comparison method and shown to be accurate and effective. Practical test results show that the CVR factor varies daily as well as seasonally: a winter day's CVR factors appear to be higher than those of a summer day. When selecting the preferred CVR feeders, the variety of CVR effects is taken into account. A Kolmogorov-Smirnov-test based probabilistic framework is used to find the probabilistic CVR performance of each feeder.

Compared with the previous efforts on evaluating CVR effects, the proposed method has four notable advantages: first, it can provide the CVR factor of any test feeder during any time; second, it does not depend on the selection of a control group or assumption of a simple linear relationship between load and its impact factors.; third, it does not require long-term CVR tests as long as the normal operation data are available for training SVR; last, it considers the varying nature of CVR effects when selecting target CVR feeders. The proposed technique can be used to analyze the CVR effects of candidate feeders before making financial investments in applying CVR.

REFERENCES

[1] D. Kirshner, "Implementation of conservation voltage reduction at commonwealth Edison," *IEEE Trans. Power Syst.*, vol. 5, pp. 1178–1182, 1990.

[2] M. Begovic, D. Novosel, B. Milosevic, and M. Kostic, "Impact of distribution efficiency on generation and voltage stability," in *Proc. 33rd Annual Hawaii Int. Conf. System Sciences, 2000*, 2000, p. 7.

[3] D. Lauria, "Conservation voltage reduction (CVR) at Northeast utilities," *IEEE Trans. Power Del.*, vol. 2, pp. 1186–1191, 1987.

[4] D. Kirshner and P. Giorsetto, "Statistical test of energy saving due to voltage reduction," *IEEE Trans. Power Apparatus Syst.*, pp. 1205–1210, 1984.

[5] B. Kennedy and R. Fletcher, "Conservation voltage reduction (CVR) at Snohomish County PUD," *IEEE Trans. Power Syst.*, vol. 6, pp. 986–998, 1991.

[6] V. Dabic, C. Siew, J. Peralta, and D. Acebedo, "BC hydro's experience on voltage VAR optimization in distribution system," in *Proc. 2010 IEEE PES*, 2010, pp. 1–7.

[7] T. L. Wilson, "Measurement and verification of distribution voltage optimization results," in *Proc. Power Energy Society General Meeting, 2010 IEEE*, 2010, pp. 1–9.

[8] A. Dwyer, R. E. Nielsen, J. Stangl, and N. S. Markushevich, "Load to voltage dependency tests at BC Hydro," *Proc. IEEE Trans. Power Syst.*, vol. 10, pp. 709–715, 1995.

[9] Z. Bashir and M. El-Hawary, "Applying wavelets to short-term load forecasting using PSO-based neural networks," *IEEE Trans. Power Syst.*, vol. 24, pp. 20–27, 2009.

[10] B.-J. Chen and M.-W. Chang, "Load forecasting using support vector machines: A study on EUNITE competition 2001," *IEEE Trans. Power Syst.*, vol. 19, pp. 1821–1830, 2004.

[11] S. Fan and L. Chen, "Short-term load forecasting based on an adaptive hybrid method," *IEEE Trans. Power Syst.*, vol. 21, pp. 392–401, 2006.

[12] E. E. Elattar, J. Goulermas, and Q. Wu, "Electric load forecasting based on locally weighted support vector regression," *IEEE Trans. Syst., Man, Cybern., C: Appl. Rev.*, vol. 40, pp. 438–447, 2010.

[13] T. Hong, P. Wang, and H. L. Willis, "A Naïve multiple linear regression benchmark for short term load forecasting," in *Proc. IEEE Power Energy Society General Meeting*, 2011, pp. 1–6.

[14] H. S. Hippert, C. E. Pedreira, and R. C. Souza, "Neural networks for short-term load forecasting: A review and evaluation," *IEEE Trans. Power Syst.*, vol. 16, pp. 44–55, 2001.

[15] R.-H. Liang and C.-C. Cheng, "Combined regression-fuzzy approach for short-term load forecasting," in *Proc. IEE Generation, Transm. Distrib.*, 2000, pp. 261–266.

[16] D. C. Park, M. El-Sharkawi, R. Marks, L. Atlas, and M. Damborg, "Electric load forecasting using an artificial neural network," *IEEE Trans. Power Syst.*, vol. 6, pp. 442–449, 1991.

[17] T. Hong and P. Wang, "On the impact of demand response: Load shedding, energy conservation, and further implications to load forecasting," in *Proc. IEEE Power Energy Society General Meeting*, 2012, pp. 1–3.

[18] E. E. Elattar, J. Y. Goulermas, and Q. Wu, "Generalized locally weighted GMDH for short term load forecasting," *IEEE Trans. Syst., Man, Cybern., C: Appl. Rev.*, vol. 42, pp. 345–356, 2012.

[19] R. B. D'Agostino and M. A. Stephens, *Goodness of Fit Techniques*. Boca Raton, FL, USA: CRC Press, 1986, vol. 68.

[20] T. Hong, "Short Term Electric Load Forecasting," Ph.D. dissertation, North Carolina State University, Raleigh, NC, USA, 2010.

[21] A. D. Papalexopoulos and T. C. Hesterberg, "A regression-based approach to short-term system load forecasting," *IEEE Trans. Power Syst.*, vol. 5, pp. 1535–1547, 1990.

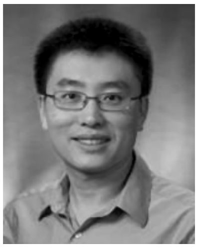
[22] V. Vapnik, *Statistical Learning Theory*. New York, NY, USA: Wiley, 1998.

Zhaoyu Wang (S'13) received the B.S. degree in electrical engineering from Shanghai Jiaotong University, Shanghai, China, in 2009, the M.S. degree in electrical engineering from Shanghai Jiaotong University in 2012, and the M.S. degree in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2012, where he is now working towards the Ph.D. degree in the School of Electrical and Computer Engineering.

His current research interests include volt/var control in power systems, conservation voltage reduction, load modeling and identification, stochastic optimization and its applications in power systems.

Miroslav M. Begovic (S'87–M'89–SM'92–F'04) is Professor and Chair of the Electric Energy Technical Interest Group in the School of Electrical Engineering at Georgia Institute of Technology. His studies focus on analysis, monitoring, and control of voltage stability and applications of phasor measurements in electrical power systems. His research is concentrated on real-time monitoring systems for control of power system dynamics, protective relaying, distribution network operation, and distributed resources in energy systems.

Dr. Begovic is a member of IEEE PES Power System Relaying Committee, a former Chair of the IEEE PES Emerging Technologies Coordinating Committee and has contributed to technical activities within IEEE and CIGRE. Dr. Begovic is a member of Sigma Xi, Tau Beta Pi, Phi Kappa Phi, and Eta Kappa Nu and currently serves as President-Elect of IEEE Power and Energy Society.



Jianhui Wang (M'07–SM'12) received the Ph.D. degree in electrical engineering from Illinois Institute of Technology, Chicago, IL, USA, in 2007.

Presently, he is a Computational Engineer with the Decision and Information Sciences Division at Argonne National Laboratory, Argonne, IL, USA.

Dr. Wang is the chair of the IEEE Power & Energy Society (PES) power system operation methods subcommittee. He is an editor of the IEEE TRANSACTIONS ON POWER SYSTEMS, the IEEE TRANSACTIONS ON SMART GRID, an Associate Editor of *Journal of Energy Engineering*, an editor of the IEEE PES Letters, and an Associated Editor of *Applied Energy*. He is also the editor of Artech House Publishers Power Engineering Book Series and the recipient of the IEEE Chicago Section 2012 Outstanding Young Engineer Award. He is also an affiliate professor at Auburn University.