Adaptively Weighted Support Vector Regression: Prognostic Application to a Historic Masonry Fort

Sez Atamturktur, M.ASCE1; Ismail Farajpour, S.M.ASCE2; Saurabh Prabhu, S.M.ASCE3; and Ashley Haydock4

Abstract: Prognostic evaluation involves constructing a prediction model based on available measurements to forecast the health state of an engineering system. One particular prognostic technique, support vector regression, has had successful applications because of its ability to compromise between fitting accuracy and model complexity in training prediction models. In civil engineering applications, compromise between fitting accuracy and model complexity depends primarily on the measured response of the system to loads other than those that are of interest for prognostic evaluation, referred to as extraneous noise in this paper. To achieve accurate prognostic evaluation in the presence of such extraneous noise, this paper presents an approach for optimally weighing fitting accuracy and complexity of a support vector regression model in an iterative manner as new measurements become available. The proposed approach is demonstrated in prognostic evaluation of the structural condition of a historic masonry coastal fortification, Fort Sumter located in Charleston, South Carolina, considering differential settlement of supports. Within this case study, the adaptive optimal weighting approach had increased forecasting accuracy over the non-weighted option. DOI: 10.1061/(ASCE)CF.1943-5509.0000517. © 2014 American Society of Civil Engineers.

Author keywords: Unreinforced masonry; Service-life assessment; Condition-based maintenance; Structural health monitoring; Historic monuments.

Introduction

In recent years, a significant amount of research has been directed towards the development of prognostic methodologies to forecast the future health state of an engineering system assisting condition-based maintenance. However, applications of these potentially useful and informative techniques to historic masonry structures are rare, if any. Developing prognostic methodologies for gradually deteriorating historic masonry monuments and infrastructure affords the possibility of ensuring structural safety, reducing maintenance costs, and preventing secondary damage of such cultural heritage.

Among available prognostic models, support vector regression (SVR) shows a distinct potential for application to historic masonry construction as it offers high accuracy, provides good generalization, and handles nonlinearity in data (Müller et al. 1997; Samanta and Nataraj 2008). The theory of SVR also recognizes that unduly complex models may have greater fitting accuracy but are less generalizable to other data sets of similar underlying processes (Myung 2000). The predictive performance of SVR then relies on having a suitable trade-off between fitting accuracy and flatness. In this paper, the term flatness refers to the opposite of model complexity. An overly complex model would fit to noise (instead of the actual trends in the data) and increase the prediction errors of the model for future data. The dual objective of SVR then seeks to find the flattest possible model while simultaneously minimizing fitting error (Smola and Schölkopf 2004). In this model training process, a user-defined weighting factor dictates the relative importance of flatness to fitting accuracy. The optimal value for this user-defined weighting factor however is dependent upon the noise resulting from not only the random errors in measurements but also the extraneous, time-varying loading conditions that are different than causal effects of long-term deterioration. For instance, if the damage in the structure is primarily caused by settlement of supports, then the variations in the structural response due to wind would be considered extraneous noise. Since such extraneous noise might vary over time, it becomes necessary to regularly adjust this weighting factor. This paper presents an approach for iteratively adjusting the weighting factor as new measurements become available to obtain a model complex enough to provide a close fit to data but simple enough to predict global trends satisfactorily. The iteratively adjusted weighted SVR approach is demonstrated on prognostic evaluation of the future health state of a historic masonry coastal fortification considering the differential settlements of the foundations. Specifically, an optimal weighting pattern is defined in which more weight is given to more recent measurements.

The article begins with a review of established literature on prognostic evaluation. Next, main concepts and governing equations for SVR are followed by a discussion on the adaptively weighted SVR approach. The historic masonry case study structure is then discussed through which adaptively weighted SVR is applied to improve forecasting accuracy in the prognostic evaluation. Finally, a discussion of the results and a summary of the contributions of this study are provided.

Note. This manuscript was submitted on April 30, 2013; approved on September 23, 2013; published online on September 25, 2013. Discussion period open until December 22, 2014; separate discussions must be submitted for individual papers. This paper is part of the Journal of Performance of Constructed Facilities. © ASCE, ISSN 0887-3828/04014057(9)/$25.00.
Background on Prognostic Evaluation of Historic Masonry

Prognosis, in the context of structural health management of engineering systems, is the estimation of a system’s remaining useful life, beyond which, corrective action is required (Saxena et al. 2009). Prognostic techniques are suitable for forecasting gradual degradation processes as opposed to damages caused by sudden unpredictable events. Thus, prognosis is an acausal problem, meaning that it requires knowledge of (or assumptions about) future loading and operating conditions to make accurate predictions. Prognosis can focus on a gradual phenomenon within an assumed loading environment; however, in such evaluation, all other loads can cause extraneous noise in the measurements (Saxena et al. 2010).

The main objective in the implementation of prognostic techniques, therefore, is to enable educated planning of maintenance of the evaluated system (Jaoude et al. 2011). Such advanced strategies have been made possible in many fields; however, prognostic evaluation of masonry heritage structures is in its infancy. With prognostic techniques fully developed and successfully applied to historic masonry monuments, timely condition-based maintenance and restoration efforts can be planned and the life of such heritage structures can be prolonged.

Masonry construction is prone to experience gradual degradations affecting structural integrity in two forms: material degradations resulting primarily from environmental impacts and structural degradations resulting primarily due to applied loads or movement of supports. Of the latter, differential support settlements are common in masonry structures because of the heavy weight of the construction and are particularly detrimental to the integrity of the structure because of the low tensile capacity of unreinforced masonry (Liu et al. 2001; Atamurtur et al. 2011). Nondestructive inspection techniques with potential to be automated that provide an indication of the global (rather than local) structural integrity are desired for prognostic evaluation of historic masonry structures. Particularly, dynamic responses, such as time-dependent accelerations or quasi-static responses, such as strains, can supply a viable solution to providing a diagnostic assessment of the structure. Other nondestructive techniques, such as the acoustic impact method (Anzani et al. 2010), the impact-echo method (Sansalone and Carino 1986), and the ultrasonic wave propagation method (Na et al. 2002), typically provide a local assessment and thus are appropriate for prognostic evaluation of a component or region of the structure.

Methodology

This section briefly discusses the theory behind support vector machine (SVM) specifically for regression and introduces an approach for adaptively weighting the model flatness to fitting accuracy in training SVR models.

Support Vector Regression

Motivated by results of the statistical learning theory (Vapnik 1998), SVM is a learning algorithm based on the structural risk-minimization principle, which finds a balance between fitting accuracy and model complexity (Xu et al. 2012). In contrast to other machine-learning approaches, such as neural networks that are prone to overfitting the data and having poor generalization capabilities, SVM can allow a predetermined degree of flatness in the model to avoid overfitting (Burges 1998; Xu et al. 2012). Furthermore, most SVMs solve a quadratic programming problem, which finds the optimal solution and assures that the obtained solution is the unique global solution.

Originally created for classification of data sets belonging to separate classes, SVM seeks to maximize the margin around the linear hyperplane dividing the linearly separable classes (Schölkopf et al. 1995; Xu et al. 2012). In cases where a linear hyperplane (i.e., model) is inappropriate for adequately separating data, a nonlinear model can be obtained by mapping the original data into a new high-dimensional feature space through the use of kernels. With the use of kernel functions, the SVM operations are performed in the input space rather than the higher dimensional feature space, thereby reducing the computational demands of high-dimensional problems (Gunn 1998).

The SVMs were extended to solve regression problems for model estimation with the addition of an appropriate cost function called the loss function (Vapnik 1998). Several types of loss functions have been offered (e.g., quadratic, ε-insensitive, and Huber) for the user to select depending on the problem (Smola and Schölkopf 2004). Among available loss functions, the quadratic loss function is proven to be a good approximation to the true loss function using Taylor series expansion (James and Stein 1961; Joseph 2004).

The basic principles of SVM for regression, known as support vector regression (SVR), can be illustrated for a training data set \([x_1, y_1], (x_2, y_2), \ldots, (x_n, y_n)\) of size \(n\). Although more complex kernel functions are available and will be discussed later, for simplicity, this discussion begins by using a linear kernel function (i.e., linear hyperplane). The linear kernel function, \(f(x)\), can be used to solve the following regression problem:

\[
f(x) = \langle w, x \rangle + bw \in \mathbb{R}^n, \quad b \in \mathbb{R}^n
\]

where \(w = \text{coefficient and } b = \text{constant offset known as bias} \). In these equations, \(\mathbb{R}^n\) refers to \(n\)-dimensional vector space over the field of real numbers. The model given in Eq. (1) is trained using a subset of the training data set that constitutes the decision boundaries or margin bounds as shown in Fig. 1 (Schölkopf et al. 1995). This subset of data points is referred to as the support vectors. The complexity of the model depends on the number of support vectors by which it is represented and is independent of the dimensionality of the input space (i.e., size of input data) (Smola and Schölkopf 2004; Drucker et al. 1997). Generally, seeking a small \(w\) in Eq. (1) decreases the percentage of data points utilized as support vectors, thus reducing model complexity and increasing model flatness (Smola and Schölkopf 2004).

The regression model is then determined by the convex optimization problem
minimize \[ \frac{1}{2} \|w\|^2 + \frac{1}{2\lambda} \sum_{i=1}^{n} (\xi_i + \xi_i^*) \]
subject to \[ \begin{cases} y_i - \langle w, x_i \rangle - b \leq \xi_i \\ \langle w, x_i \rangle + b - y_i \leq \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \] (2)

in which the regularization parameter \( \lambda \) is traditionally a prespecified constant that determines the effect of the slack parameters; \( \xi, \xi^* \) (i.e., the errors calculated by the loss function) on the objective function. When \( \lambda \to 0 \), maximizing fitting accuracy (i.e., minimizing fitting error) is the main objective of the optimization. Conversely, when \( \lambda \to \infty \), maximizing model flatness (i.e., minimizing complexity) becomes the main objective of the optimization. Therefore, applying \( \lambda > 0 \) achieves a compromise between fitting accuracy and flatness.

By minimizing Eq. (2), a balance is found between complexity, \( 1/2 \times \|w\|^2 \), and overall fitting loss, \( 1/2\lambda \times \sum_{i=1}^{n} (\xi_i + \xi_i^*) \). With the selection of an appropriate \( \lambda \), this balance ensures that the obtained model generalizes well preventing the model from fitting to noise (also known as overfitting). Therefore, the model sensitivity to noise can be reduced with the selection of a suitable \( \lambda \) for a given data set.

The loss function used in this study is the quadratic loss function, \( L_{quad} \), which can be written as follows:

\[ L_{quad}[f(x) - y] = [f(x) - y]^2 \] (3)

To measure the error between the observed values \( y \) and estimated outputs \( f(x) \) for a given input \( x \), Eq. (3) uses the conventional least squares error criterion as shown in Fig. 2.

The solution to Eq. (2) in the quadratic loss function formulation is given by

\[ \max_{\alpha, \alpha^*} W(\alpha, \alpha^*) = \max_{\alpha, \alpha^*} -\frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)(x_i, x_j) \\
+ \sum_{i=1}^{l} (\alpha_i - \alpha_i^*)y_i - \lambda \sum_{i=1}^{l} [\alpha_i^2 + (\alpha_i^*)^2] \] (4)

where \( \alpha, \alpha^* \) represent the Lagrange multipliers. By exploiting the Karush-Kuhn-Tucker conditions (Karush 1939)

\[ \bar{\alpha}_i, \bar{\alpha}_i^* = 0, \quad i = 1, \ldots, l \] (5)

the optimization problem can be simplified as

\[ \min_{\beta} \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} \beta_i \beta_j (x_i, x_j) - \sum_{i=1}^{l} \beta_i y_i + \lambda \sum_{i=1}^{l} \beta_i^2 \] (6)

with constraints

\[ \sum_{i=1}^{l} \beta_i = 0 \] (7)

where \( \beta \) represents the Lagrange multipliers. The regression model is given by Eq. (1) where

\[ \bar{\omega} = \sum_{i=1}^{l} \beta_i x_i \quad \bar{b} = -\frac{1}{2} \langle \bar{\omega}, (x_r + x_s) \rangle \] (8)

In Eqs. (4), (6), and (8), the dot product, \( \langle x_r, x_s \rangle \), can be replaced by a kernel function to map the linear SVR formulation to solve a nonlinear problem, a process widely known as nonlinear mapping (Gunn 1998). Various kernel functions, such as polynomial, spline, and radial basis functions are available for nonlinear mapping. Because of their flexibility and consistency of fitting and predicting with minimal residual error in comparison to other kernels, splines are a common kernel function of choice in SVR modeling (Mammen 1997; Rajasekaran et al. 2008); thus, the remainder of the manuscript will focus on the spline kernel.

**Adaptively Weighted Support Vector Regression**

The trade-off between fitting accuracy and flatness of an SVR greatly affects the predictive performance of the prognostic evaluation. This principle is evident in Fig. 3. Models that are too simple, as shown by \( \lambda = 2 \) in Fig. 3(a), may neither be able to fit the available data nor be able to generalize the trends well. Models that are too complex in contrast may accurately fit the available data, as shown by \( \lambda \to 0 \) in Fig. 3(a), but may not be able to generalize the trends well. Therefore, there is an optimal degree of flatness, as shown by \( \lambda = 0.01 \) in Fig. 3(b) that finds a suitable compromise between fitting error and model flatness. This optimal model flatness depends both on the characteristics of the inherent
nonlinearities of the underlying degradation process and on the extraneous noise present in the measurements. Thus, the optimal $\lambda$ is that which generalizes global trends in the presence of noise.

In Fig. 4, the role of extraneous noise when selecting $\lambda$ is demonstrated focusing on goodness of fit to a given trend. In noise-free data sets, $\lambda \to 0$ (i.e., giving zero weight to flatness) may provide a model that fits the trend well as shown in Fig. 4(a). As noise increases, however, a larger $\lambda$ is required, meaning that more weight must be given to flatness than fitting accuracy, to achieve a similar fitting to the trend as presented in Figs. 4(b and c). Therefore, $\lambda$ must be correctly determined for a given data set to ensure reliable predictions of the future health state of the system. A similar study could have been conducted for examining the effect of the nonlinearity of the model on $\lambda$ selection.

In the published literature, cross-validation has been used for selecting $\lambda$ by utilizing hold-out experiments; however, this technique focuses solely on fitting accuracy (in an interpolative manner) rather than prediction accuracy (in an extrapolative manner) (Stone 1974; Jaakkola and Haussler 1999; Smola and Schölkopf 2004). Because a prognostic evaluation requires accurate extrapolative projections of the future health of the structure, the focus in this paper is to improve the forecasting accuracy of the model rather than its closeness of fit to available data. Hence, in this paper, the optimal $\lambda$ is selected by that which predicts with the least $L_1$ norm prediction error (i.e., summation of absolute values of errors) over the hold-out set is chosen as the optimal $\lambda$. During the forecasting stage, this optimal $\lambda$ is then used to train a refined model using the total available data set (i.e., up to $n$) to predict the forecasting set (i.e., from $n$ to $p$). The adaptively weighted approach then repeats this process as additional measurements become available by adding new measurements to the training set and updating $\lambda$ accordingly.

Algorithm 1. Detailed steps of this process are shown in the pseudocode

Begin

Input SVR parameters

$X =$ independent variable

$Y =$ dependent variable

$h =$ number of hold-out points

$f =$ number of forecasting points

$P =$ total number of iterations

$m =$ index of final point in preliminary training set

$n =$ index of final point in hold-out set

$p =$ index of final point in forecasting set

For $i = 1$ to $P$

For $\lambda = 10^{-15}$ to $10^{5}$

Train a support vector regression model (Gunn 1998) using preliminary training set, $X_1$ to $X_m$, and forecast the hold-out set, $X_{m+1}$ to $X_{n}$. Compute the $L_1$ norm residual error of the predicted hold-out set by comparison to the corresponding subset of $Y$.

End

Choose optimal $\lambda$ as that which gave the least prediction error of the hold-out set.

The basic steps of this adaptively weighted prognostic approach can be demonstrated on an initial data set of $n$ points. In Fig. 5, the data set is divided into three parts: the preliminary training set consisting of the first $m$ points, the hold-out set consisting of the following $h$ points, and the forecasting set consisting of the next $f$ points. During the preliminary stage, optimal $\lambda$ is selected. For this, multiple candidate $\lambda$ values are tested in their ability to predict the hold-out set of $h$ points from $m$ to $n$, where $n = m + h$. The candidate $\lambda$ producing the model with the least $L_1$ norm prediction error (i.e., summation of absolute values of errors) over the hold-out set is chosen as the optimal $\lambda$. During the forecasting stage, this optimal $\lambda$ is then used to train a refined model using the total available data set (i.e., up to $n$) to predict the forecasting set (i.e., from $n$ to $p$). The adaptively weighted approach then repeats this process as additional measurements become available by adding new measurements to the training set and updating $\lambda$ accordingly.
Train a support vector regression model using training set, \( X_1 \) to \( X_m \), and predict the forecasting set, \( X_{n+1} \) to \( X_p \), where \( X_p = X_{(n+f)} \).
Compute the residual error of the predicted forecasting set.
Define new input parameters: \( X_{m+1} = X_n \), \( X_{n+1} = X_p \).
End.
End.

Case Study

Coastal fortifications built as defense mechanisms in protecting important seaports and harbors were once the cornerstone of national defense in the United States (McGovern and Smith 2006). Today, these coastal fortifications, many of which are over 150 years old, are considered structures of national heritage. Over their lifetime, these structures are subject to harsh coastal environmental and operational conditions leading to material and structural degradations.

To successfully preserve these historically important edifices for future generations, timely maintenance is imperative. Prognostic evaluation can assure such timely maintenance campaigns.

Fort Sumter, in Charleston, South Carolina, where the first shots of the American Civil War were fired in 1861 (National Park Service 1984) is one such historically important fort that is in need of accurate structural assessment and prognostic evaluation. There is evidence that differential settlement of the foundation has been occurring at Fort Sumter leading to extensive cracks throughout the masonry casemates. Thus, the weighted SVR prognostic technique is applied to predict the future changes in the measurable response features for one of the casemates of Fort Sumter considering gradual settlement of foundations.

Case Study Structure: Fort Sumter National Monument

The construction of the pentagonal-shaped clay masonry fort began in 1829 on an artificial island. In the years of the Civil War, Fort Sumter witnessed several battles that severely damaged the structure (National Park Service 1984). After several rounds of demolition and reconstruction, Fort Sumter was declared a national monument in 1948. The fort has since been maintained by the National Park Service and is currently accessible to visitors [Fig. 6 (Holbrook Field Trips 2013)].

Finite Element Model Development

The finite element (FE) model of the single casemate used in this study as shown in Fig. 7 is developed in Ansys 13.0 by incorporating data from on-site inspections and evaluations discussed in detail in Atamturktur and Prabhu (2013). Material properties are obtained through laboratory tests conducted on core samples of the masonry and a masonry prism specimen from fallen debris. The three-dimensional (3D) laser scanning is performed to obtain the precise as-is geometry of the casemate with which the FE model is constructed preserving key geometrical features. The FE model is developed using SOLID65 elements that are specialized for modeling concrete-like brittle materials (Ozen 2006; Mahini et al. 2007; Li and Atamturktur 2013). The SOLID65 element uses a smeared crack analogy to account for deformations because of cracking and crushing of the material. The linear material properties of the model are calibrated to experimentally obtained modal parameters (i.e., first two natural frequencies and mode shapes) (see Atamturktur and Sevim 2012 for a detailed discussion on model calibration).

Because the barrel-vaulted casemates are built adjacent to but detached from the scarp wall, the scarp wall and the casemate are two independent structural entities. Therefore, contact elements that allow sliding and separation (but do not allow penetration) of two adjacent components are used to model this interface. A dynamic hammer impact test was used to calibrate the friction coefficient accounting for the friction and cohesion (if any) at the interface to represent this possible sliding action in the FE model. To take into consideration the lateral interaction with the adjacent casemates, adjacent casemates are represented using substructuring techniques. To keep the size of the model to a manageable level, the foundations of the casemate are idealized as a series of linear springs having finite stiffness. The reader is invited to read Atamturktur and Prabhu (2013) for further details of the model development process.

Simulations of Support Settlement

The FE model used to simulate support settlement is shown in Fig. 8, where the casemate of interest is the center casemate with the adjacent casemates modeled as substructures. The ground beneath the casemates can be visualized as a rectangular plane as shown in Fig. 8(a). By tilting this rectangular plane in the direction perpendicular to the external wall as shown in Fig. 8(b), the settlement configuration representing settlement of the external wall is simulated. The ground plane of the casemate is simulated to...
gradually settle with a maximum displacement ($\Delta$) under the scarp wall of from 0 to 100 mm at increments of 2.5 mm.

Many different global response features can be measured as part of a health monitoring framework (see Prabhu and Atamturktur 2012 for a review of these features). In this paper, the strain at the two control point locations, Point 1 and Point 2, shown in Fig. 9, are monitored during settlement simulations. A more extensive prognostic campaign may involve mounting a larger number of sensors distributed throughout the structure. The location of the measurement points must be selected based not only on the engineering judgment regarding critical stress concentrations but also on the accessibility of the points for sensor placement. As shown in Fig. 9, Point 1 is located at the base of the pier, and Point 2 is located at the springing of the arch (at the top of the pier). In practical applications, methods for optimal sensor placement, as explained in Kammer (1991) and Prabhu and Atamturktur (2013), can also be implemented to maximize the information gain from measurements.

Next, randomly generated noise is added to the strain response at a magnitude of 10% of the maximum strain according to the following equations:

$$y_{\text{noise}} = y_{\text{sim}} + \text{rand} \times y_{\text{sim}};$$

$$\text{rand} \in R,$$

$$\sum_{i=1}^{n} \text{rand}_i = 0,$$

$$\text{STD(} \text{rand}_1, \text{rand}_2, \ldots, \text{rand}_n) = 10\%.$$  

where $y_{\text{noise}}$ represents that simulated data after adding the noise values; $y_{\text{sim}}$ represents the original simulated data; rand is a real random value with the mean value of the zero and standard deviation of the 10%. The resulting first principal strains at the two control points, Point 1 and Point 2, are plotted in Fig. 10.

**Prognostic Evaluation Using Adaptively Weighted SVR**

In the prognostic evaluation, 15 data points simulating the strain response of the casemate under settlement up to 40 mm, at which first cracks start to appear, are used. To determine the initial $\lambda$ value, the first 10 of these data points are used in the preliminary training set (which corresponds to 27.5 mm settlement) and the next five data points are used as the hold-out set (from 27.5 to 40 mm settlement) (refer back to Fig. 5). Although in this paper, the selection of the number of data points in the preliminary and hold-out sets are somewhat arbitrary, the rule of thumb is to leave a larger number of data points in the training set than in the hold-out set.

Multiple candidate $\lambda$ values between $10^{-15}$ and $10^5$ (10 $\lambda$ values for each multiple of 10 from $10^{-15}$ to $10^5$) are tested to find the optimal $\lambda$ that yields the minimal error in predicting the hold-out set. With the identified optimal $\lambda$, a refined SVR model is trained and is executed to forecast the next five data points (from

---

**Fig. 8.** (a) Initial model configuration on level surface; (b) settlement configuration

**Fig. 9.** Locations Point 1 and Point 2 of monitored strains during settlement (circled)

**Fig. 10.** Settlement-induced strains obtained from FE model of (a) Point 1; (b) Point 2 with added noise
40 to 52.5 mm settlement). This process is repeated as new measurement data become available, and the optimal $\lambda$ is updated during each iteration. In this case study, a total of five iterations are completed to reach 100 mm settlement, at which the geometric deformations of the structure become visually detectable making an automated sensing system to detect and monitor the structural behavior unnecessary. Thus, the optimal $\lambda$ is updated four times after it is initially determined in the first trial. The predicted response, prediction error, and adaptively refined optimal $\lambda$ obtained as a result of this analysis are displayed in Fig. 11 for Point 1 and Fig. 12 for Point 2 [results shown after the vertical dashed line in Figs. 11 and 12(a and b) are the compiled results of the five forecasting iterations]. For comparison, the predicted response and prediction error of an SVR model trained using a constant $\lambda \rightarrow 0$, which gives all weight to fitting error and none to flatness, are also included in the figures.

As evidenced in Figs. 11 and 12, the adaptively weighted SVR predicts the settlement-induced strains with less than half as much error as the nonweighted approach (Table 1). The noise added to the simulated data is nonstationary in nature. The results are repeatable, however, for various forms of noise. Therefore, the distinct advantage of the adaptive approach is its ability to recover the optimal $\lambda$.

<table>
<thead>
<tr>
<th>SVR approach</th>
<th>Point 1 strain: unitless</th>
<th>Point 2 strain: unitless</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adaptively weighted</td>
<td>0.0719</td>
<td>0.0057</td>
</tr>
<tr>
<td>Nonweighted</td>
<td>0.1898</td>
<td>0.0178</td>
</tr>
</tbody>
</table>

Table 1. Total Prediction Error for Adaptively Weighted SVR and Nonweighted SVR
as extraneous noise fluctuations occur over time, as is the case in practical in situ monitoring applications.

Furthermore, the previous analysis is repeated 20 times each with different randomly generated noise added to the original simulated data sets. Figs. 13 and 14 show the forecasts (past the dashed line) for Point 1 and for Point 2, respectively. The adaptively weighted SVR shows greatly reduced uncertainty (i.e., less than half the standard deviation) in forecasts in comparison to the nonweighted option as evident in Table 2, where the average of the standard deviation of the prediction error at each data point is displayed.

**Conclusion**

Although SVR is known for its superior prognostic abilities, the performance of this machine-learning technique is reliant on the selection of an appropriate regularization parameter, $\lambda$, determining the trade-off between fitting accuracy and model complexity (i.e., flatness). The optimal trade-off is greatly affected by the presence of time-variant extraneous noise within measurements, which is common during in situ monitoring applications. Therefore, an ideal process for selecting optimal $\lambda$ is one in which the model sensitivity to noise is decreased.

Within this paper, an adaptive weighting approach for SVR is developed, which first determines the optimal $\lambda$ based on forecasting accuracy and then uses this optimal $\lambda$ to develop a refined model for future predictions. As additional data become available in time, the optimal $\lambda$ is updated allowing the new model to be adjusted for fluctuations in noise intensity. Thus, the most suitable model complexity for a given data set is selected for each set of predictions.

In testing the performance of this approach on the simulated settlement response of one of the casemates of Fort Sumter, the adaptively weighted SVR showed greatly increased forecasting accuracy in forecasts over the nonweighted approach. Furthermore, the adaptively weighted SVR approach also showed decreased uncertainty when 20 realizations of extraneous noise were evaluated.

The developed adaptively weighted SVR has potential to be incorporated in a structural health monitoring process to ultimately assist in preserving the cultural heritage by predicting its future structural integrity. However, future direction in research should focus on determination of appropriate damage sensitive features and corresponding monitoring techniques for prognosis of historic masonry structures. Furthermore, a failure threshold indicating the structure’s end of life must also be defined. Such a threshold can only be defined by developing a link between nondestructive measurements and the remaining load-carrying capacity of the structure (as proposed in Atamturktur et al. 2012), which is the

| Table 2. Average Standard Deviation of Prediction Error for Adaptively Weighted SVR and Nonweighted SVR |
|---------------------------------------------------------------|---------------------------------------------------------------|
| SVR approach    | Point 1 strain: unitless | Point 2 strain: unitless |
| Adaptively weighted | 0.0016                     | 0.0002                     |
| Nonweighted      | 0.0041                     | 0.0004                     |

primary attribute of concern in prognostic evaluation, thus warranting future work.

Acknowledgments
This work was sponsored in part by the PTT Grants program of the National Center for Preservation Technology and Training (NCPTT) of the Department of Interior: the Grant Agreement Number MT-2210-11-NC-02.

References

Ansys 13.0 [Computer software]. Ansys, Inc., Pittsburgh, PA.