Article

Reconstruction to Sensor Measurements Based on a Correlation Model of Monitoring Data

Wei Lu [1,*], Jun Teng [1,2], Chao Li [1] and Yan Cui [1]

[1] Department of Civil and Environmental Engineering, Harbin Institute of Technology Shenzhen Graduate School, Shenzhen 518055, China; tengj@hit.edu.cn (J.T.); lc2013.hit@gmail.com (C.L.); cui.yan@stu.hit.edu.cn (Y.C.)

[2] College of Civil Engineering, Fujian University of Technology, Fuzhou 350118, China

* Correspondence: lu.wei@hit.edu.cn; Tel: +86-755-2603-3056

Abstract: A sensor failure will lead to sensor measurement distortion, and may reduce the reliability of the whole structure analysis. This paper studies the method of monitoring information reconstruction based on the correlation degree. For the faulty sensor, the correlation degree of the normal response of this sensor and the measurements of the other sensors is calculated, which is also called the correlation degree of reconstructed variables and response variables. By comparing the correlation degrees, the response variables, which are needed to establish the correlation model, are determined. The correlation model between the reconstructed variables and the response variables is established by the partial least square method. The value of the correlation degrees between the reconstructed variables and the response variables, the amount of the monitoring data which is used to determine the coefficients of the correlation model, and the number of the response variables are used to discuss the influence factors of the reconstruction error. The stress measurements of structural health monitoring system of Shenzhen Bay Stadium is taken as an example, and the effectiveness of the method is verified and the practicability of the method is illustrated.

Keywords: structural health monitoring; responses reconstruction; correlation model; stress measurements; Shenzhen Bay Stadium

1. Introduction

The sensors are the basic instruments of the structural health monitoring system, while the measurements of sensors are the necessary information for estimating the structural safety. However, the operation duration of a sensor is limited and the sensor measurements may be abnormal because of sensor failure or environmental influences [1]. The technique for reconstructing the sensor information should be developed as a result, which can ensure saving the complete monitoring information for structural safety estimation. The monitoring information is diverse in the different measured locations and the different types of structural responses, which are correlated with each other to some extent [2]. That is to say, it is necessary to reconstruct the information for the failure sensor using the effective correlation model between the failure sensor and the other normal sensors.

For the sensor information reconstruction, the researchers proposed several methods to solve the question. For the problem of data distortion in the structural health monitoring system of a bridge, Huang [3] analyzed the relationship of the monitoring data by using Granger causality testing [4] in which the sensor data with a closer relationship is selected to be the input vector for an extreme learning machine to recover the missing sensor signal data. However, the individual prediction value could be a large error, and the training results need to be further improved because the tendency of the predicted data is not consistent with the theory value. The radial basis function (RBF) neural

network model is another method to recover the unreliable data, which is also easy to fall into a local optimum [5]. The Kalman filtering method was proposed to build the model between structural displacements and structural accelerations [6], in which the structural model parameters and external excitation should be known in advance. The method of data compression sampling is used for data recovery of wireless sensors, while the recovery data error would be larger for the accelerations when the sparse of the original data becomes worse [7]. Feng [8] proposed a data recovery method based on a least square support vector machine. A forecasting model based on the method is established, and signal recovery is accomplished by on-line learning diagnosis. However, this method mainly aims at the data reconstruction of the sensor within short duration, and the effect is not good for a high sampling rate and for long-term failure prediction.

This paper, thus, focuses on the structural responses and the correlation model between the reconstructed variable and the response variable is established, which can be used for reconstructing the sensor information. In this paper, the correlation degree and the partial least square method are applied to the data reconstruction of the fault sensors. As the strain sensors are placed on the roof steel structure of Shenzhen Bay Stadium, the strain measurements represent the distributed responses of the structure and have an inherent association, which is suitable to verify the effectiveness of the proposed method in the paper as well.

2. Correlation Model of Monitoring Data

2.1. Correlation Analysis of Monitoring Data

The monitoring structural responses from different locations and different types of sensors are in the correlated relationship. The Pearson correlation coefficient is an index to measure the linear correlation degree of structural responses.

The structural response matrix of m measuring points under the load in n time steps is denoted as $\mathbf{X}$:

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & ... & \mathbf{x}_m \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix} \tag{1}$$

$$\mathbf{x}_i = \begin{bmatrix} x_{i1} & x_{i2} & \cdots & x_{in} \end{bmatrix}^T \tag{2}$$

In the Equations (1) and (2), $\mathbf{x}_i$ represents the measurements of the measuring point i under the load of n time steps. When the fault sensor works in normal, the measurement of the measuring point where the fault sensor located is denoted as $\mathbf{y}$. The Pearson product-moment correlation coefficient [9] of the measurement from the arbitrary measuring point and the measuring point where the faulty sensor located is denoted as ρ_i

$$\rho_i = \frac{\text{cov}(\mathbf{x}_i, \mathbf{y})}{\sigma_{\mathbf{x}_i} \sigma_{\mathbf{y}}} = \frac{E(\mathbf{x}_i - \mu_{\mathbf{x}_i})(\mathbf{y} - \mu_{\mathbf{y}})}{\sigma_{\mathbf{x}_i} \sigma_{\mathbf{y}}} \tag{3}$$

$$\sigma_{\mathbf{x}_i} = \frac{1}{n-1} \sum_{k=1}^{n} \left(\mathbf{x}_i(k) - \mu_{\mathbf{x}_i} \right)^2 \tag{4}$$

$$\sigma_{\mathbf{y}} = \frac{1}{n-1} \sum_{k=1}^{n} \left(\mathbf{y}(k) - \mu_{\mathbf{y}} \right)^2 \tag{5}$$

$$\text{cov}(\mathbf{x}_i, \mathbf{y}) = \frac{1}{n-1} \sum_{k=1}^{n} \left(\mathbf{x}_i(k) - \mu_{\mathbf{x}_i} \right) \left(\mathbf{y}(k) - \mu_{\mathbf{y}} \right) \tag{6}$$

In Equation (3), $\sigma_{\mathbf{x}_i}$ and $\sigma_{\mathbf{y}}$ represent the response variance of the measurement point i and the fault sensor, respectively; $\text{cov}(\mathbf{x}_i, \mathbf{y})$ is covariance, $\mathbf{x}_i(k)$ and $\mathbf{y}(k)$ represent the measurement of point i and the fault sensor in the kth time step, respectively; $\mu_{\mathbf{x}_i}$ and $\mu_{\mathbf{y}}$ are the mean value of the measurements $\mathbf{x}_i$ and $\mathbf{y}$, respectively.

2.2. Correlation Modeling Using Partial Least Squares and Multiple Regression

2.2.1. Reconstructed Variable and Response Variables

The correlation coefficients of the normal measurement of fault sensor and the measurements of the other sensors are calculated, which are also called the correlation coefficients of the reconstructed variable and response variables. The correlation matrix can be denoted as **P**:

$$\mathbf{P} = \begin{bmatrix} \rho_1 & \rho_2 & \cdots & \rho_m \end{bmatrix} \tag{7}$$

The correlation coefficients are sorted from larger value to small value based on their absolute values, while the updated correlation coefficient's sorting matrix can be denoted as $\mathbf{D}_P$:

$$\mathbf{D}_P = \{(\rho)_1 \quad \cdots (\rho)_i \quad \cdots \quad (\rho)_m \} \tag{8}$$

$(\rho)_i$ represents the ith correlation coefficient absolute value, while $(\rho)_1$ is the largest correlation coefficient absolute value and $(\rho)_m$ is the smallest correlation coefficient absolute value, respectively. The corresponding measurement to $(\rho)_i$ is denoted as $(\mathbf{x})_i$, while the updated structural response sorting matrix can be denoted as $\mathbf{D}_X$:

$$\mathbf{D}_X = \{(\mathbf{x})_1 \quad \cdots (\mathbf{x})_i \quad \cdots \quad (\mathbf{x})_m \} \tag{9}$$

The faulty sensor information is considered as the reconstructed variable $\mathbf{Y}_P$, and p measurements are selected as response variables $\mathbf{X}_P$:

$$\mathbf{X}_P = \{(\mathbf{x})_1 \quad (\mathbf{x})_2 \quad \cdots \quad (\mathbf{x})_p \} \tag{10}$$

$$\mathbf{Y}_P = \{\mathbf{y}\} \tag{11}$$

2.2.2. Establishment of Correlation Model Using the Partial Least Square Method

The reconstructed variables and response variables are standardized as:

$$(\mathbf{x})_i^* = \frac{(\mathbf{x})_i - \mu_{(\mathbf{x})_i}}{\sigma_{(\mathbf{x})_i}} \tag{12}$$

$$\mathbf{y}^* = \frac{\mathbf{y} - \mu_{\mathbf{y}}}{\sigma_{\mathbf{y}}} \tag{13}$$

The response variables $\mathbf{X}_P$ and the reconstructed variable $\mathbf{Y}_P$ are standardized into $\mathbf{E_0}$ and $\mathbf{F_0}$, respectively:

$$\mathbf{E_0} = \{(\mathbf{x})_1^* \quad (\mathbf{x})_2^* \quad \cdots \quad (\mathbf{x})_p^* \} \tag{14}$$

$$\mathbf{F_0} = \{\mathbf{y}^*\} \tag{15}$$

The main components set of $\mathbf{E_0}$ and $\mathbf{F_0}$ can be obtained by using principal component analysis, which are denoted as **T** and **U**, respectively:

$$\mathbf{T} = \left\{ t_1 \quad t_2 \quad \ldots \quad t_h \quad \ldots \quad t_A \right\} \tag{16}$$

$$\mathbf{U} = \left\{ \begin{array}{ccccc} u_1 & u_2 & ... & u_h & ... & u_A \end{array} \right\} \tag{17}$$

The regression equations regarding $\mathbf{E_0}$ and $\mathbf{F_0}$ using partial least squares [10] and first principal components are:

$$\mathbf{E_0} = t_1 \mathbf{p_1}^T + \mathbf{E_1} \tag{18}$$

$$\mathbf{F_0} = t_1 \mathbf{r_1}^T + \mathbf{F_1} \tag{19}$$

where the regression coefficient vector is:

$$\mathbf{p_1} = \frac{\mathbf{E}_0^T t_1}{\|t_1\|^2} \tag{20}$$

$$\mathbf{r_1} = \frac{\mathbf{F}_0^T t_1}{\|t_1\|^2} \tag{21}$$

The regression equations for the hth principal component are:

$$\mathbf{E}_{h-1} = t_h \mathbf{p}_h^T + \mathbf{E}_h \tag{22}$$

$$\mathbf{F}_{h-1} = t_h \mathbf{r}_h^T + \mathbf{F}_h \tag{23}$$

where the regression coefficient vector is:

$$\mathbf{p}_h = \frac{\mathbf{E}_{h-1}^T t_h}{\|t_h\|^2} \tag{24}$$

$$\mathbf{r}_h = \frac{\mathbf{F}_{h-1}^T t_h}{\|t_h\|^2} \tag{25}$$

The relationship between the hth principal component t_h and the hth axis of $\mathbf{E_0}$ is:

$$t_h = \mathbf{E}_{h-1} \mathbf{w}_h \tag{26}$$

where $\mathbf{w}_h$ is the unit eigenvector corresponding to the maximum eigenvalue $\theta_h{}^2$ of matrix $\mathbf{E}_{h-1}{}^T \mathbf{F}_{h-1} \mathbf{F}_{h-1}{}^T \mathbf{E}_{h-1}$.

According to the Equations (18), (19), (22), and (23), $\mathbf{E_0}$ and $\mathbf{F_0}$ can be expressed as:

$$\mathbf{E_0} = t_1 \mathbf{p_1}^T + t_2 \mathbf{p_2}^T + \cdots + t_h \mathbf{p}_h^T + \mathbf{E}_h \tag{27}$$

$$\mathbf{F_0} = t_1 \mathbf{r_1}^T + t_2 \mathbf{r_2}^T + \cdots + t_h \mathbf{r}_h^T + \mathbf{F}_h \tag{28}$$

Substituting Equation (26) into Equation (22) yields:

$$\mathbf{E}_h = \mathbf{E}_{h-1}(\mathbf{I} - \mathbf{w}_h \mathbf{p}_h^T) \tag{29}$$

Substituting Equations (26) and (29) into Equation (27) yields:

$$\mathbf{E}_h = \mathbf{E_0} \prod_{j=1}^{h} (\mathbf{I} - \mathbf{w}_j \mathbf{p}_j^T) \tag{30}$$

Substituting Equation (30) into Equation (26) yields:

$$t_h = \mathbf{E}_{h-1} \mathbf{w}_h = \mathbf{E_0} \prod_{j=1}^{h-1} (\mathbf{I} - \mathbf{w}_j \mathbf{p}_j^T) \mathbf{w}_h = \mathbf{E_0} \mathbf{w}_h^* \tag{31}$$

$$\mathbf{w}_h* = \prod_{j=1}^{h-1}(\mathbf{I} - \mathbf{w}_j\mathbf{p}_j{}^T)\mathbf{w}_h \tag{32}$$

Substituting Equation (31) into Equation (28) yields:

$$\mathbf{F}_0 = \mathbf{E}_0\mathbf{w}_1{}^*\mathbf{r}_1{}^T + \cdots + \mathbf{E}_0\mathbf{w}_h{}^*\mathbf{r}_h{}^T + \mathbf{F}_h \tag{33}$$

Assuming the number of the principal components is h when the principal components satisfy the cross-validation principle. Meanwhile, $\mathbf{F}_h$ approaches 0, which can be ignored. Equation (33) is written as:

$$\mathbf{F}_0 = \mathbf{E}_0(\mathbf{w}_1{}^*\mathbf{r}_1{}^T + \cdots + \mathbf{w}_h{}^*\mathbf{r}_h{}^T) \tag{34}$$

Substituting Equations (14) and (15) into Equation (34) yields:

$$\mathbf{y}* = [(\mathbf{x})_1^* \ldots (\mathbf{x})_i^* \ldots (\mathbf{x})_p^*]
\begin{bmatrix}
\sum\limits_{j=1}^{h} \mathbf{r}_j\mathbf{w}_{j1}{}^* \\
\ldots \\
\sum\limits_{j=1}^{h} \mathbf{r}_j\mathbf{w}_{ji}{}^* \\
\ldots \\
\sum\limits_{j=1}^{h} \mathbf{r}_j\mathbf{w}_{jp}{}^*
\end{bmatrix} \tag{35}$$

If:

$$\alpha_i{}^* = \sum_{j=1}^{h} \mathbf{r}_j\mathbf{w}_{ji}{}^* \tag{36}$$

then:

$$\mathbf{y}^* = \alpha_1{}^*(\mathbf{x})_1{}^* + \cdots + \alpha_p{}^*(\mathbf{x})_p{}^* \tag{37}$$

Substituting Equations (12) and (13) into Equation (37), the correlation model between the reconstructed variable and the response variables is:

$$\mathbf{y}(i) = \alpha_1(\mathbf{x}(i))_1 + \cdots + \alpha_p(\mathbf{x}(i))_p + \alpha_0 \tag{38}$$

2.2.3. The Reconstruction Error Evaluation Index

The average absolute error between the reconstructed value and actual value e_a is:

$$e_a = \frac{1}{n}\sum_{i=k}^{n}|\mathbf{y}(k) - \mathbf{y}_0(k)| \tag{39}$$

The average relative error between the reconstructed value and actual value e_b is:

$$e_b = \frac{1}{n}\sum_{k=1}^{n}\left|\frac{\mathbf{y}(k) - \mathbf{y}_0(k)}{\mathbf{y}_0(k)}\right| \times 100\% \tag{40}$$

where $\mathbf{y}(k)$ represents the reconstructed response of kth time step of the supposed fault sensor, and $\mathbf{y}_0(k)$ represents the actual response of kth time step of the supposed fault sensor, respectively.

3. Reconstruction to Stress Measurements of Shenzhen Bay Stadium

3.1. Project Description

The structural health monitoring system of Shenzhen Bay Stadium began to be implemented in 2010, while the continuous monitoring of the steel roof in both the construction phase and service phase began to be carried out at the same time [11]. The stress measurements of the sensors located at the support members are chosen as the basic analysis data for example. The layout of sensors located at the support members are shown in Figures 1 and 2. There are six supports, two rods of which are being monitored and four stress monitoring points are arranged on each rod; the sensor installation of support is shown in Figure 3. For example, the number of eight stress monitoring points at support 3-1 can be denoted as 3-1-1, 3-1-2, 3-1-3, 3-1-4, 3-1-5, 3-1-6, 3-1-7, and 3-1-8, respectively.
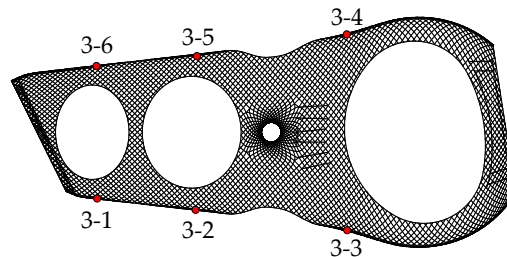

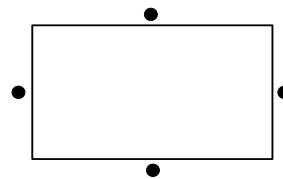**Figure 1.** Stress monitoring point of the support connecting rod.

Figure 2. Distribution of the stress sensor.

Figure 3. The sensor installation of two support connecting rods.

3.2. The Correlation Model of the Reconstructed Variable and the Response Variables

The sensor at location 3-1-4 is assumed to be failure on 16 May 2014. The stress monitoring information of this point is considered as the reconstructed variable. The measurements of the other locations are considered as the available response variables. The stress monitoring data from 7 May 2014 to 10 May 2014 are selected to calculate the correlation coefficients between measurements from location 3-1-4 and the other locations, which are shown in Table 1.

Table 1. The correlation degree of stress monitoring data between stress sensor 3-1-4 and other stress sensors.

Ranking	Sensor Location	Correlation Coefficient	Ranking	Sensor Location	Correlation Coefficient
1	3-1-8	0.6831	18	3-1-2	0.2076
2	3-6-6	0.5437	19	3-6-1	0.1975
3	3-3-2	0.5194	20	3-5-1	0.1962
4	3-2-8	0.4656	21	3-1-3	0.1623
5	3-1-7	0.4212	22	3-2-7	0.1113
6	3-2-4	0.4121	23	3-6-8	0.0915
7	3-1-6	0.4006	24	3-3-7	0.0742
8	3-2-6	0.3735	25	3-2-2	0.0632
9	3-5-7	0.3688	26	3-3-1	0.0624
10	3-2-3	0.3677	27	3-3-3	0.0547
11	3-5-4	0.3608	28	3-2-5	0.0424
12	3-6-3	0.3512	29	3-5-2	0.0214
13	3-5-5	0.3425	30	3-5-8	0.0156
14	3-3-6	0.3289	31	3-5-3	0.0141
15	3-5-6	0.2745	32	3-6-4	0.0108
16	3-1-5	0.2608	33	3-2-1	0.0031
17	3-1-1	0.2487	-	-	-

In Equation (38), there is no exact number demanded of the response variables, which may depend on the detailed examples and the special value of the correlation coefficients. Generally speaking, the historical measurements can be used to identify the effectiveness when the different number of responses variables selected. If four response variables are chosen to build the correlation model, that is to say, $p = 4$, then:

$$\mathbf{D}_P = \{0.6831 \quad -0.5437 \quad -0.5194 \quad 0.4656 \} \tag{41}$$

$$\begin{aligned} \mathbf{X}_{P1} &= \left\{ (\mathbf{x})_1 \quad (\mathbf{x})_2 \quad (\mathbf{x})_3 \quad (\mathbf{x})_4 \right\} \\ &= \left\{ \sigma_{3-1-8}(t) \quad \sigma_{3-6-6}(t) \quad \sigma_{3-3-2}(t) \quad \sigma_{3-2-8}(t) \right\} \end{aligned} \tag{42}$$

$$\mathbf{Y}_{P1} = \{\mathbf{y}_1\} = \{\sigma_{3-1-4}(t)\} \tag{43}$$

The correlation model of the reconstructed variable and the response variable is:

$$\mathbf{y}_1(i) = 1.01(\mathbf{x}(i))_1 + 0.05(\mathbf{x}(i))_2 - 0.09(\mathbf{x}(i))_3 + 0.10(\mathbf{x}(i))_4 - 20.73 \tag{44}$$

The stress monitoring data of the stress sensor 3-6-6, 3-3-2, 3-1-8, and 3-2-8 are used to reconstruct the stress measurements of the sensor located at point 3-1-4 on 16 May 2014. The contrast curves between reconstructed values and actual values are shown in Figure 4. The reconstructed values have the same trend as the actual values, and there are small differences between them. The average absolute error is 0.89 MPa and the relative error is 2.71%.
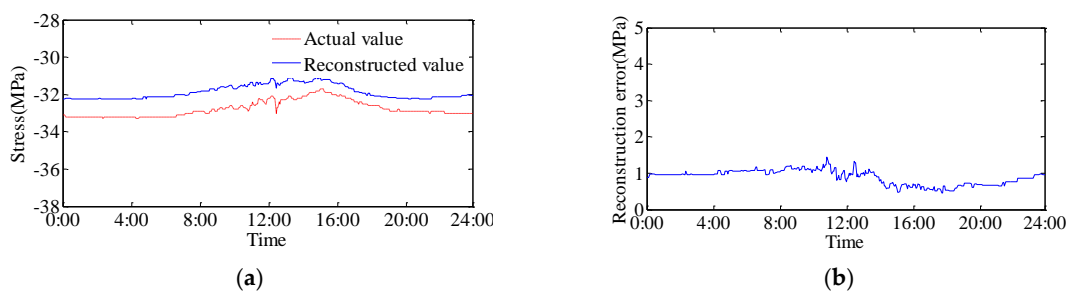

Figure 4. The comparison between the reconstructed value and the actual value: (**a**) the curve of reconstructed value and the actual value; and (**b**) the curve of absolute error.

3.3. Influence of Correlation Degree on Reconstruction Error

In order to discuss the influences of the selection of response variables, two additional cases are discussed, which are:

$$\mathbf{X}_{P2} = \left\{ \begin{array}{cccc} (\mathbf{x})_{13} & (\mathbf{x})_{14} & (\mathbf{x})_{15} & (\mathbf{x})_{16} \end{array} \right\} \\ = \left\{ \begin{array}{cccc} \sigma_{3-5-5}(t) & \sigma_{3-3-6}(t) & \sigma_{3-5-6}(t) & \sigma_{3-1-5}(t) \end{array} \right\} \tag{45}$$

and:

$$\mathbf{X}_{P3} = \left\{ \begin{array}{cccc} (\mathbf{x})_{30} & (\mathbf{x})_{31} & (\mathbf{x})_{32} & (\mathbf{x})_{33} \end{array} \right\} \\ = \left\{ \begin{array}{cccc} \sigma_{3-5-8}(t) & \sigma_{3-5-3}(t) & \sigma_{3-6-4}(t) & \sigma_{3-2-1}(t) \end{array} \right\} \tag{46}$$

The correlation model using $\mathbf{X}_{P2}$ is:

$$\mathbf{y}_2(i) = 0.02(\mathbf{x}(i))_{13} + 0.01(\mathbf{x}(i))_{14} + 0.39(\mathbf{x}(i))_{15} + 0.02(\mathbf{x}(i))_{16} - 23.85 \tag{47}$$

The contrast curves between reconstructed values and actual values are shown in Figure 5. The average absolute error is 0.97 MPa and the relative error is 2.97%.
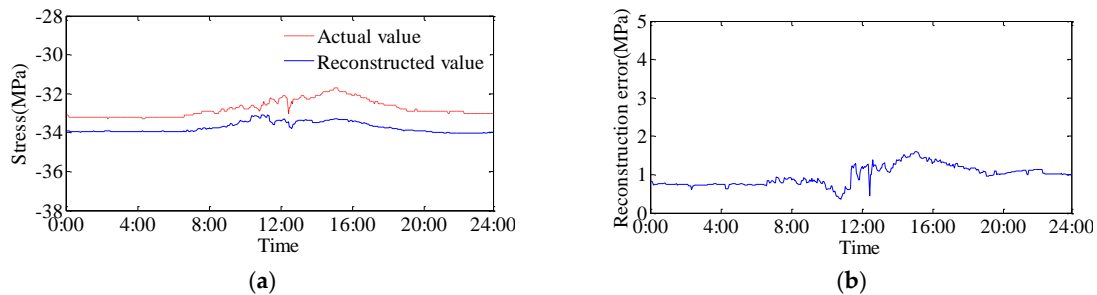

Figure 5. The reconstructed value and the actual value using response variable $\mathbf{X}_{P2}$: (**a**) the curve of the reconstructed value and the actual value; and (**b**) the curve of reconstruction error.

The correlation model using $\mathbf{X}_{P3}$ is:

$$\mathbf{y}_3(i) = -0.01(\mathbf{x}(i))_{30} + 0.01(\mathbf{x}(i))_{31} - 0.01(\mathbf{x}(i))_{32} - 0.01(\mathbf{x}(i))_{33} - 32.03 \tag{48}$$

The contrast curves between reconstructed values and actual values are shown in Figure 6. The average absolute error is 1.03 MPa and the relative error is 3.13%.
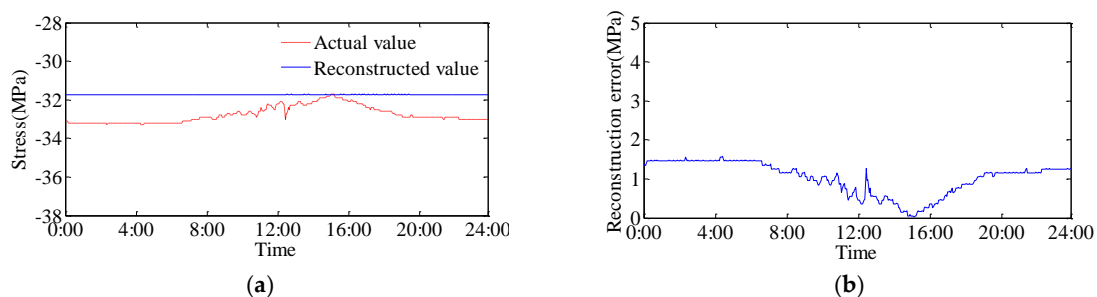

Figure 6. The reconstructed value and the actual value using response variable $\mathbf{X}_{P3}$: (**a**) the curve of the reconstructed value and the actual value; and (**b**) the curve of reconstruction error.

The errors of the reconstructed information using different correlation coefficients of the response variables are summarized in Table 2.

Table 2. The summarized errors using different response variables.

Response Variables	e_a (MPa)	e_b
$\mathbf{X}_{P1}$	0.89	2.71%
$\mathbf{X}_{P2}$	0.97	2.97%
$\mathbf{X}_{P3}$	1.03	3.13%

It can be seen from Table 2 that the errors become larger when the coefficients of response variables become smaller, in which the average absolute error and relative error are minimum when $\mathbf{X}_{P1}$ is chosen as the response variables. It indicated that the response variables with closer correlation to the reconstructed variable should be chosen, which is better for the reconstruction of the sensor measurements.

3.4. Influence of Monitoring Data Quantity on Reconstruction Error

In order to discuss the influences of the selection of response variables, two additional cases are discussed. The original stress monitoring data is collected from 7 May 2014 to 10 May 2014 to calculate the correlation coefficients between measurements from location 3-1-4 and the other locations. The stress monitoring data of two additional cases are from 7 May 2014 to 12 May 2014 and from 7 May 2014 to 14 May 2014, respectively.

The correlation model using stress monitoring data from 7 May 2014 to 12 May 2014 is:

$$\mathbf{y}_4(i) = 0.89(\mathbf{x}(i))_1 + 0.09(\mathbf{x}(i))_2 - 0.20(\mathbf{x}(i))_3 + 0.04(\mathbf{x}(i))_4 - 33.62 \tag{49}$$

The contrast curves between reconstructed values and actual values are shown in Figure 7. The average absolute error is 0.65 MPa and the relative error is 1.99%.
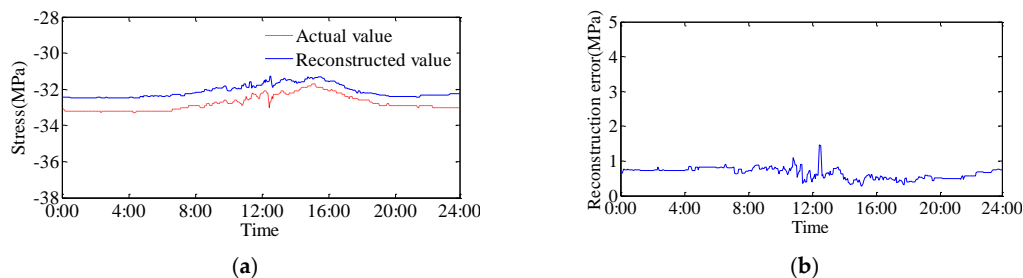

Figure 7. The reconstructed value and the actual value using data from 7 May 2014 to 12 May 2014: (**a**) the curve of reconstructed value and the actual value; and (**b**) the curve of reconstruction error.

The correlation model using stress monitoring data from 7 May 2014 to 14 May 2014 is:

$$\mathbf{y}_5(i) = 0.35(\mathbf{x}(i))_1 - 0.16(\mathbf{x}(i))_2 - 0.11(\mathbf{x}(i))_3 + 0.03(\mathbf{x}(i))_4 - 39.45 \tag{50}$$

The contrast curves between the reconstructed values and the actual values are shown in Figure 8. The average absolute error is 0.22 MPa and the relative error is 0.68%.

The errors of the reconstructed information using stress monitoring data from different time periods are summarized in Table 3.

It can be seen from Table 3 that the errors become smaller when the quantity of stress monitoring data become larger, in which the average absolute error and relative error are a minimum when the stress monitoring data from 7 May 2014 to 12 May 2014 is chosen to build the correlation model. This indicated that the larger quantity of stress monitoring data should be used, which is better for reconstruction of sensor measurements. In theory, the number of measurements larger than the unknown coefficients is the basic requirement for regression function identification; in fact, there are

noises and temperature effects included in the measurements, the measurements in at least one full day are better.

Table 3. The summarized errors using stress monitoring data from different time periods.

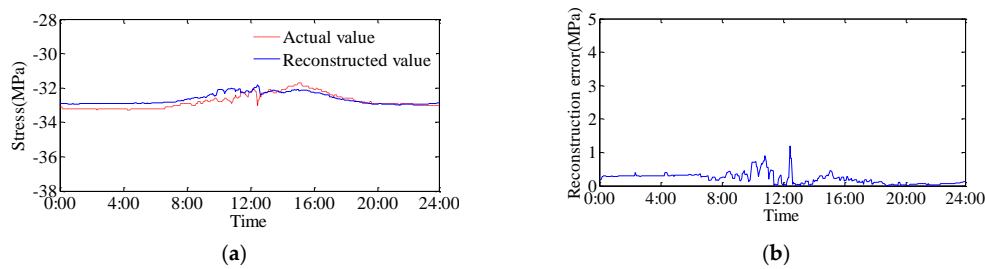| Data Acquisition Time	e_a (MPa)	e_b
2014.05.07~2014.05.10	0.89	2.71%
2014.05.07~2014.05.12	0.65	1.99%
2014.05.07~2014.05.14	0.22	0.68%

(a)　　　　　　　　　　　　　　(b)

Figure 8. The reconstructed value and actual value using data from 7 May 2014 to 14 May 2014: (**a**) the curve of the reconstructed value and the actual value; and (**b**) the curve of the reconstruction error.

3.5. Influence of Response Variable Number on Reconstruction Error

In order to discuss the influences of the different number of response variables, three additional cases are discussed, which are:

$$\mathbf{X}_{P5} = \left\{ (\mathbf{x})_1 \quad (\mathbf{x})_2 \quad (\mathbf{x})_3 \quad (\mathbf{x})_4 \quad (\mathbf{x})_5 \right\} \tag{51}$$

$$\mathbf{X}_{P6} = \left\{ (\mathbf{x})_1 \quad (\mathbf{x})_2 \quad (\mathbf{x})_3 \quad (\mathbf{x})_4 \quad (\mathbf{x})_5 \quad (\mathbf{x})_6 \right\} \tag{52}$$

$$\mathbf{X}_{P7} = \left\{ (\mathbf{x})_1 \quad (\mathbf{x})_2 \quad (\mathbf{x})_3 \quad (\mathbf{x})_4 \quad (\mathbf{x})_5 \quad (\mathbf{x})_6 \quad (\mathbf{x})_7 \right\} \tag{53}$$

The correlation model using $\mathbf{X}_{P5}$ is:

$$\mathbf{y}_6(i) = 1.33(\mathbf{x}(i))_1 + 0.10(\mathbf{x}(i))_2 - 0.03(\mathbf{x}(i))_3 - 0.08(\mathbf{x}(i))_4 - 0.20(\mathbf{x}(i))_5 - 32.82 \tag{54}$$

The contrast curves between reconstructed values and actual values are shown in Figure 9. The average absolute error is 0.80 MPa and the relative error is 2.45%.
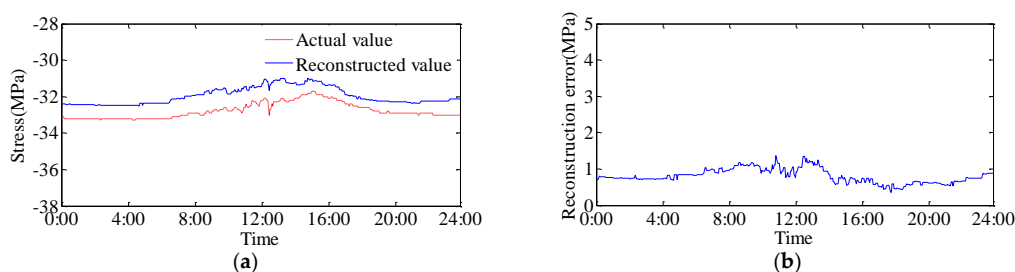

(a)　　　　　　　　　　　　　　(b)

Figure 9. The reconstructed value and actual value using response variables $\mathbf{X}_{P5}$: (**a**) the curve of the reconstructed value and the actual value; and (**b**) the curve of reconstruction error.

The correlation model using $\mathbf{X}_{P6}$ is:

$$\mathbf{y}_7(i) = 1.33(\mathbf{x}(i))_1 + 0.11(\mathbf{x}(i))_2 - 0.04(\mathbf{x}(i))_3 - 0.11(\mathbf{x}(i))_4 - 0.16(\mathbf{x}(i))_5 + 0.08(\mathbf{x}(i))_6 - 30.95 \tag{55}$$

The contrast curves between reconstructed values and actual values are shown in Figure 10. The average absolute error is 0.80 MPa and the relative error is 2.44%.
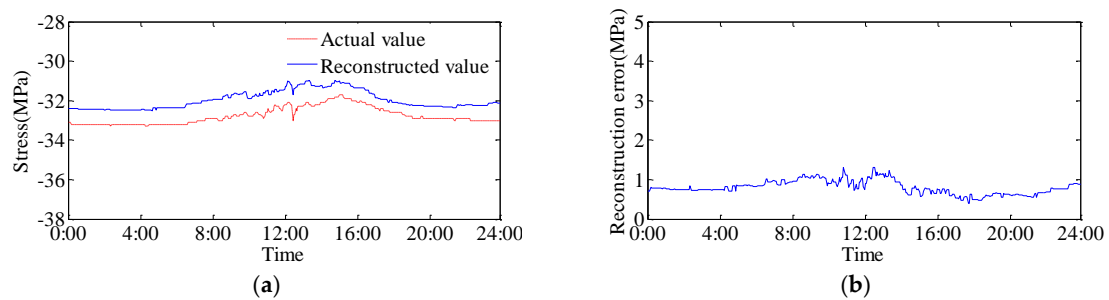

Figure 10. The reconstructed value and actual value using response variables $\mathbf{X}_{P6}$: (**a**) the curve of the reconstructed value and the actual value; and (**b**) the curve of reconstruction error.

The correlation model using $\mathbf{X}_{P7}$ is

$$\mathbf{y}_8 = 1.30(\mathbf{x})_1 + 0.09(\mathbf{x})_2 - 0.07(\mathbf{x})_3 - 0.08(\mathbf{x})_4 - 0.19(\mathbf{x})_5 + 0.12(\mathbf{x})_6 + 0.13(\mathbf{x})_7 - 26.93 \tag{56}$$

The contrast curves between the reconstructed values and the actual values are shown in Figure 11. The average absolute error is 0.77 MPa and the relative error is 2.36%.
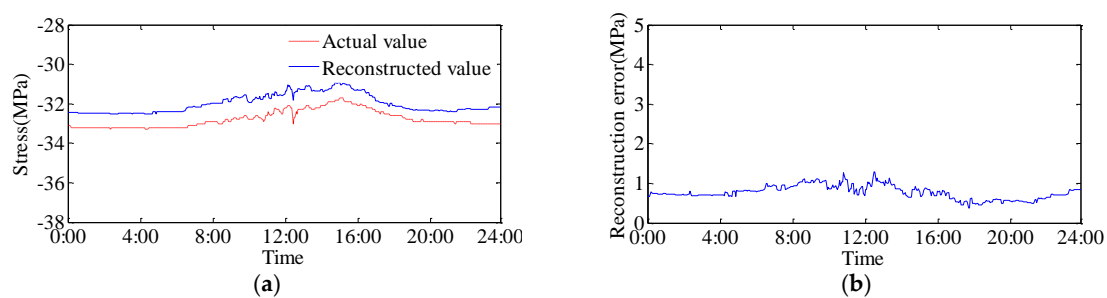

Figure 11. The reconstructed value and actual value of the response variables $\mathbf{X}_{P7}$: (**a**) the curve of reconstructed value and the actual value; and (**b**) the curve of the reconstruction error.

The errors of the reconstructed information using different number of the response variables are summarized in Table 4.

Table 4. The summarized errors using different number of the response variables.

Response Variable	e_a (MPa)	e_b
$\mathbf{X}_{P1}$	0.89	2.71%
$\mathbf{X}_{P5}$	0.80	2.45%
$\mathbf{X}_{P6}$	0.80	2.44%
$\mathbf{X}_{P7}$	0.77	2.36%

It can be seen from Table 4 that the errors become smaller when the number of response variables become larger, in which the average absolute error and relative error are minimum when $\mathbf{X}_{P7}$ is chosen as the response variables. However, there is only a slight difference between these discussed cases, and the number of the response variables with the same level of correlation coefficients influencing the errors of the reconstructed variables.

4. Conclusions

In this paper, a method of sensor information reconstruction based on the correlation degree is proposed. The monitoring information of a faulty sensor can be reconstructed by the correlation model and the measurements of other normal sensors. Based on the structural health monitoring system of Shenzhen Bay Stadium, the stress measurements of sensors located at support members are used to verify the proposed method. Furthermore, the influences for the errors of reconstructed monitoring information of a faulty sensor caused by three kinds of factors are discussed in the paper, which are the selection of response variables with different correlation, the quantity of the measurements from different time periods, and the number of response variables. The selection of response variables and the quantity of the measurements from different time periods are two main factors, which influence the reconstructed information errors for the faulty sensor directly. The number of the response variables influences the errors of reconstructed information a little if the selected response variables are in the same correlation level to reconstructed variable. Furthermore, the proposed method is verified its effectiveness to a real world project.

Author Contributions: Wei Lu wrote the paper; Wei Lu and Jun Teng conceived and designed the Structural Health Monitoring (SHM) System of Shenzhen Bay Stadium; Chao Li conducted experimental data analysis; Yan Cui conducted experimental data acquisition and collation.

Conflicts of Interest: The authors declare no conflict of interest

References

1. Yeung, W.T.; Smith, J.W. Damage detection in bridges using neural networks for pattern recognition of vibration signatures. *Eng. Struct.* **2005**, *27*, 685–698. [CrossRef]
2. Agrawal, R.; Imielinski, T.; Swami, A.N. Mining association rules between sets of items in large databases. In Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data, 26–28 May 1993; ACM: Washington, DC, USA; pp. 207–216.
3. Huang, Y.W.; Wu, D.G.; Li, J. Structural healthy monitoring data recovery based on extreme learning machine. *Comput. Eng.* **2011**, *37*, 241–243. (In Chinese)
4. Hu, S.R.; Chen, W.M.; Cai, X.X.; Liang, Z.B.; Zhang, P. Research of deflection data compensation by applying artificial neural networks. *In Proceedings of the 2nd International Conference on Structural Health Monitoring of Intelligent Infrastructure, Shenzhen, China, 16–18 November 2005*; Taylor & Francis-Balkema: Leiden, The Netherlands; pp. 755–758.
5. Huang, G.B.; Zhu, Q.Y.; Siew, C.K. Extreme learning machine: Theory and applications. *Neurocomputing* **2006**, *70*, 489–501. [CrossRef]
6. Ma, T.W.; Bell, M.; Lu, W.; Xu, N.S. Recovering structural displacements and velocities from acceleration measurements. *Smart. Struct. Syst.* **2014**, *14*, 191–207. [CrossRef]
7. Bao, Y.Q.; Li, H.; Sun, X.D.; Yu, Y.; Ou, J.P. Compressive sampling-based data loss recovery for wireless sensor networks used in civil structural health monitoring. *Struct. Health Monit.* **2013**, *12*, 78–95. [CrossRef]
8. Feng, Z.G.; Shida, K.; Wang, Q. Sensor fault detection and data recovery based on LS-SVM predictor. *Chin. J. Sci. Instrum.* **2007**, *28*, 193–197.
9. Pearson, K. Notes on regression and inheritance in the case of two parents. *Proc. R. Soc. Lond.* **1895**, *58*, 240–242. [CrossRef]
10. Tenenhaus, M.; Esposito Vinzi, V.; Chatelinc, Y.-M.; Lauro, C. PLS path modeling. *Comput. Stat. Data Anal.* **2005**, *48*, 159–205. [CrossRef]
11. Teng, J.; Lu, W.; Cui, Y.; Zhang, R.G. Temperature and displacement monitoring to steel roof construction of Shenzhen Bay Stadium. *Int. J. Struct. Stab. Dyn.* **2016**, *16*, 1640020. [CrossRef]