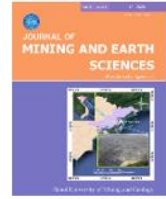




Journal of Mining and Earth Sciences

Website: <http://jmes.humg.edu.vn>



Application of correlation and regression analysis between GPS - RTK and environmental data in processing the monitoring data of cable - stayed bridge



Tinh Duc Le ^{1,*}, Hien Van Le ², Linh Thuy Nguyen ², Thanh Kim Thi Nguyen ¹, Duy Tien Le ³

¹ Faculty of Geomatics and Land Administration, Hanoi University of Mining and Geology, Vietnam

² University of Transport and Communications, Hanoi, Vietnam

³ The branch of Hanoi University of Natural Resources and Environment in Thanh Hoa Province, Vietnam

ARTICLE INFO

Article history:

Received 28th Sept. 2020

Revised 29th Nov. 2020

Accepted 31st Dec. 2020

Keywords:

Cable - stayed bridge,
Correlation analysis,
GPS - RTK,
Monitoring,
Regression analysis,
Structural health.

ABSTRACT

Structural Health Monitoring system - SHMs has been playing a vital role in monitoring large - scale structures during their performance in a lifetime, especially with the long - span bridge, such as a suspended bridge or cable - stayed bridge. In a SHM system, many kinds of sensors are used to set up at the specific locations in order to monitor and detect any changes of structures in real - time based on the changes of monitoring data as well as the changes of correlation among monitoring data types. This paper proposes a method of applying the correlation and regression analysis for processing the displacement monitoring data acquired by GPS - RTK considering the effects of environmental factors such as temperature and wind - speed. The results show that the air - temperature has high correlation with the displacements of a cable - stayed bridge acquired by GPS - RTK measurement along to specific directions while the wind - speed has low correlation. Then the general displacement of the target bridge could be recognized and regression equation is also built to predict the bridge displacement under effects of the air temperature.

Copyright © 2020 Hanoi University of Mining and Geology. All rights reserved.

1. Introduction

Structural Health Monitoring (SHM) has been

using successfully to monitor the super - structures during their operations, such as high - rise buildings and long - span structures. In a SHM system, there are many kinds of sensors setting on target structures for observing different objectives, such as capturing dynamic or static structural responses by using strain sensor, stress

*Corresponding author

E - mail: leductinhdtct@gmail.com

DOI: 10.46326/JMES.2020.61(6).07

sensor or accelerometer etc; monitoring environmental factors by using temperature and wind - speed sensors (Kaloop and Li, 2009).

For large - scale structures, monitoring the deformation of structures is an important task that can assess the structural health and then detect any damage to structures. Long - span bridges such as cable - stayed bridges or suspended - supported bridges have two kinds of deformation: long - term and short - time deformation. Long - term deformation is often caused by environmental factors, while short - time deformation is mainly caused by dynamic inputs, such as wind, earthquake, traffic, etc. (Kaloop and Li, 2009; 2011; Celebi, 2000).

Using the interferometer or some electronic distance measuring instruments is helpful to monitor the displacements of a structure. Although these methods can provide high accuracy results, they still have some shortcomings in the application. They neither cannot measure the large displacements of structures, especially with long - span bridge; nor measure in real - time or in inconvenient weather condition, etc., (Cheng and Zheng, 2002). Recently, the Global Navigation Satellite System (GNSS) has been using to monitor the displacement of a super - structure in an SHM system, especially in SHM of a long span bridge, such as Stonecutters bridge in Hong Kong, Akashi Kaikyo bridge in Japan, Ting Kau cable - stayed bridge in Hong Kong, etc. In Vietnam, GPS technology has been used in some cable - stayed bridges, such as Can Tho bridge, Tran Thi Ly bridge, Nhat Tan bridge and Bach Dang bridge. GPS is considered a high - cost method in SHM system. However, it has many advantages, such as it is less affected by weather condition; it can measure the displacements of a specific point in 3D dimension at a millimeter level of accuracy (Kaloop and Li, 2009; Cheng and Zheng, 2002).

Considering the long - term monitoring of structures, data processing is vital for recognizing the structural changes during their operation. Some studies showed that environmental factors significantly affect the long - term monitoring data (Sohn et al., 1999; Cornwell et al., 1999; Farrar et al., 2000). The correlation analysis method is often used to analyze long - term monitoring data, recognizing the effects of environmental or

operational factors on the outcome displacement data. High correlated coefficients of any factors show strong influence to the outcome displacement data (Cornwell et al., 1999; Farrar et al., 2000; Omenzetter and Brownjohn, 2005; Omenzetter and Brownjohn, 2006; Sohn et al., 2000; Hien And Mayuko, 2015; Hien et al., 2015). Besides, a regression algorithm is an effective method to use in analyzing time - series data to detect outlier and use for further prediction. To defining a regression model is a fitted model of a given time - series data by assessing the determination coefficient and testing the fitted redundant between model and data (Sanford Weisberg, 2005; Shumway and Stoffer, 2010; Peter and Annick, 1987).

This study analyzes the long - term monitoring displacements of a real cable - stayed bridge acquired by GPS - RTK measurement considering the effects of environmental factors such as air - temperature and wind - speed. Time - series monitoring data of the target cable - stayed bridge was acquired for analysis, including GPS displacements, air - temperature, and wind - speed. Correlation analysis was then adopted to figure out how the air - temperature and wind - speed effect GPS displacement data, from which the global deformation of the target bridge could be recognized in some significant directions. The regression model in both mono variant and multivariant variables was used to describe displacement modeling of the target bridge. Results of regression models were then used to assess which environmental factor and which significant direction of the target bridge is useful for analyzing the structural changes.

2. Correlation analysis

The correlation among variables can be analyzed using two kinds of the formula: single correlation and multiple correlations.

2.1. Single correlation analysis

Assume $\{X_i, Y_i\}$ ($i=1 \div n$) are two random variables; the correlation coefficient r_{XY} between variable X and variable Y can be calculated by the following steps:

- Step 1: Calculating the correlation coefficient between X and Y:

$$r_{XY} = \frac{\frac{\sum_i (X_i - \bar{X})(Y_i - \bar{Y})}{n}}{\sqrt{\frac{\sum_i (X_i - \bar{X})^2}{n}} \sqrt{\frac{\sum_i (Y_i - \bar{Y})^2}{n}}} = \frac{\bar{X}\bar{Y} - \bar{X}\bar{Y}}{\sqrt{\bar{X}^2 - (\bar{X})^2} \sqrt{\bar{Y}^2 - (\bar{Y})^2}} \quad (1)$$

Where:

$$\begin{cases} \bar{X} = \frac{\sum_i X_i}{n}; \bar{Y} = \frac{\sum_i Y_i}{n}; \bar{XY} = \frac{\sum_i X_i Y_i}{n} \\ \bar{X}^2 = \frac{\sum_i X_i^2}{n}; \bar{Y}^2 = \frac{\sum_i Y_i^2}{n} \end{cases} \quad (2)$$

The correlation coefficient calculating by equation (1) shows the relationship between two variables X and Y , which has a value domain from -1 to +1. If coefficient r_{XY} is closed to +1 or -1, it means that variable X and Y have a very high correlation. In the contrary, if r_{XY} is closed to 0, it means variable X and Y have a very low correlation.

- Step 2: Assessing the stability of correlation coefficient that depends on the time interval of monitoring:

1 - With a large number of times of monitoring ($n \geq 50$):

$$\sigma_r \approx \frac{1 - r^2}{\sqrt{n}} \quad (3)$$

Then, the correlation between X and Y satisfies the condition follows:

$$|r| \geq 3\sigma_r \quad (4)$$

2 - If $n < 50$, the Fisher equation is used:

$$Z = \frac{1}{2} \ln \frac{1+r}{1-r} \quad (5)$$

Variance of Z can be calculated by:

$$\sigma_r \approx \frac{1}{\sqrt{n-3}} \quad (6)$$

and checking the correlation condition by

$$|Z| \geq 3\sigma_Z \quad (7)$$

Figure 1 describes the correlation between two variables X and Y :

The correlation coefficient can be considered a “effect coefficient” when the correlation coefficient is approximately +1 or -1, which means the effect between two variables is very high.

95% confidence interval of the correlation coefficient: the correlation coefficient is affected by the oscillation of variables. Thus, it is necessary to calculate a 95% confidence interval of the correlation coefficient. To calculate the 95% confidence interval of the correlation coefficient, we have to use the standard deviation of the correlation coefficient calculated by:

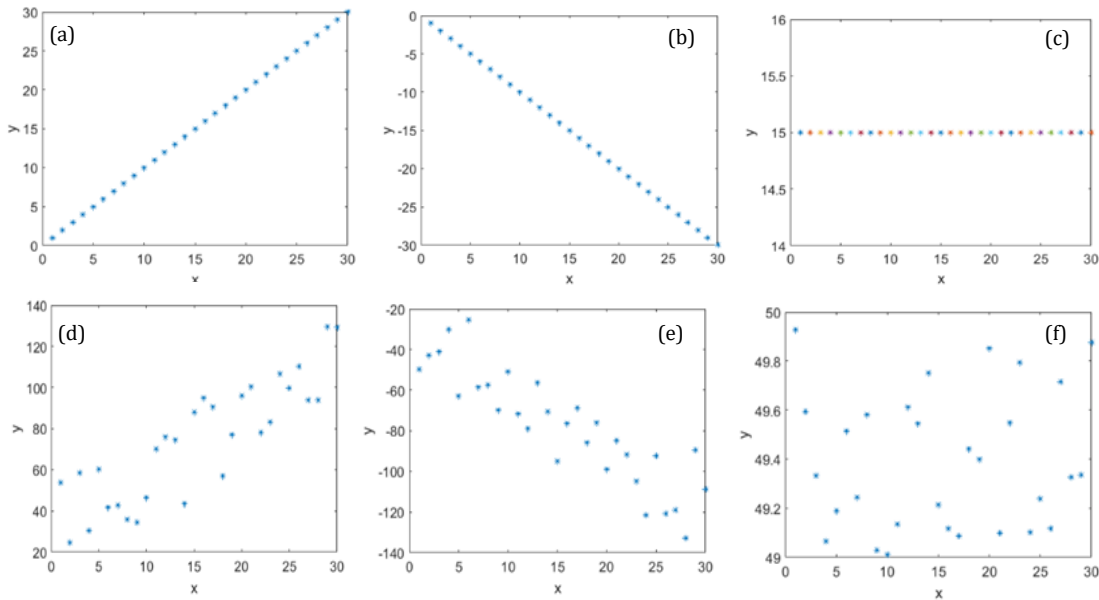


Figure 1. Examples of correlation between two variables.
(a) $r = 1$; (b) $r = -1$; (c) $r = 0$; (d) $r = 0.86$; (e) $r = -0.88$; (f) $r = 0$.

$$s_r = \frac{\sqrt{1-r^2}}{\sqrt{n-2}} \quad (8)$$

Equation 8 shows that s_r and r are dependent; thus, using an unbiased method is necessary. Ronald A. Fisher showed that calculating s_r of a function of r is an impartial method. By this calculation, substitution variable z can be defined by:

$$z = \frac{1}{2} \log \frac{1+r}{1-r} \quad (9)$$

Then, the standard deviation of z is calculated by:

$$s_z = \frac{1}{\sqrt{n-3}} \quad (10)$$

Hence, the 95% confidence interval of z can be substituted to the correlation coefficient by:

$$r = \frac{e^{2z} - 1}{e^{2z} + 1} \quad (11)$$

2.2. Multiple correlation analysis

Considering p random quantities X_1, X_2, \dots, X_p , which are measured independently in n times described in Table 1.

Table 1. A sample of monitoring data.

Period	Random quantities					
	X_1	X_2	...	X_k	...	X_p
1	X_{11}	X_{21}	...	X_{k1}	...	X_{p1}
2	X_{12}	X_{22}	...	X_{k2}	...	X_{p2}
...
i	X_{1i}	X_{2i}	...	X_{ki}	...	X_{pi}
...
n	X_{1n}	X_{2n}	...	X_{kn}	...	X_{pn}

A random variable is specified by expectation $M(x_i)$, variance, and correlating moment K_{ij} :

$$K = \{K_{ij}\}_{j=1, p}^{i=1, p} \quad (12)$$

Estimation of expectation, variance, and correlating moment can be defined by (11):

$$M[x_k] = \frac{1}{n} \sum_{i=1}^n x_{ki}, (k = 1, 2, \dots, p) \quad (13)$$

$$D_{x_k} = \frac{1}{n-1} \sum_{i=1}^n (x_{ki} - M[x_k])^2 \quad (14)$$

$$K_{ki} = \frac{1}{n-1} \sum_{i=1}^n (x_{ki} - M[x_k])(x_{ii} - M[x_i]) \quad (15)$$

Dividing the correlating matrix (15) to the corresponding variance $\sigma_k = \sqrt{D}$ and $\sigma_i = \sqrt{D}$, the correlating matrix can be defined as:

$$r = \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1k} \\ r_{21} & r_{22} & \dots & r_{2k} \\ \dots & \dots & \dots & \dots \\ r_{k1} & r_{k2} & \dots & r_{kk} \end{pmatrix} \quad (16)$$

Analyzing correlation between p random quantities (X_i, X_j, X_k) , the dependence between 2 quantities can be determined by partial correlation coefficients (Khanh Tran and Quang Phuc Nguyen, 2010; Khanh Tran and Duc Tinh Le, 2010; Duc Tinh Le, 2012), calculated by the equation below:

$$r_{12,34\dots p} = \frac{r_{12,34\dots(p-1)} - r_{1p,34\dots(p-1)}r_{2p,34\dots(p-1)}}{\sqrt{(1-r_{1p,34\dots(p-1)}^2)(1-r_{2p,34\dots(p-1)}^2)}} \quad (17)$$

Statistical assessment of correlation coefficients is done following Fisher criterion (assume: analyze four random quantities):

$$F_\Phi = \frac{R_{1,234}^2(n-m)}{(1-R_i^2)(m-1)} \geq F_q \quad (18)$$

where n is the number of quantities, m is the number of parameters. If condition (18) is correct, then the correlation coefficient R_i is accepted.

3. Regression Establishment

3.1. Establishment of mono variant regression

The mono variant regression is used to describe the correlation between two variables X and Y , as shown in the equation below:

$$Y = a.X + b \quad (19)$$

Parameters a, b are determined by the least square principle applying for n measurement couple (Y, X) , which are:

$$\{(Y_i, X_i)\} = \{(Y_1, X_1), (Y_2, X_2), \dots, (Y_n, X_n)\},$$

then set of equations can be written as below (Khanh Tran, Duc Tinh Le, 2010):

$$\begin{cases} [X^2]a + [X]b - [XY] = 0 \\ [X]a + nb - [Y] = 0 \end{cases} \quad (20)$$

Solving the set equations combining with equation (1), parameters a and b are then defined:

$$a = r_{XY} \frac{\sqrt{X^2 - (\bar{X})^2}}{\sqrt{Y^2 - (\bar{Y})^2}} \quad (21)$$

$$b = \bar{Y} - a\bar{X}$$

3.2. Establishment of multivariant regression

Regression function $f(X_2, X_3, \dots, X_p)$ describes linear dependence between variable Y and p variables (X_1, X_2, \dots, X_p) following the least square principle, as shown:

$$\begin{aligned} (Y - f(X_2, X_3, \dots, X_p))^2 \\ = E(Y - f(X))^2 = \min E \end{aligned} \quad (22)$$

When $p > 2$, the multivariant equation is:

$$\begin{aligned} Y = f(X_1, X_2, \dots, X_p) \\ = a_0 + a_1X_1 + \dots + a_pX_p \end{aligned} \quad (23)$$

Notation:

$$\begin{cases} A = \begin{pmatrix} 1 & x_{11} & \dots & x_{p1} \\ 1 & x_{12} & \dots & x_{p2} \\ \dots & \dots & \dots & \dots \\ 1 & x_{1n} & \dots & x_{pn} \end{pmatrix} \\ L = \begin{pmatrix} Y_1 \\ Y_2 \\ \dots \\ Y_n \end{pmatrix}; Z = \begin{pmatrix} a_0 \\ a_1 \\ \dots \\ a_p \end{pmatrix} \end{cases} \quad (24)$$

According to equation (22) and condition (23), the matrix form of the set equation can be established as (Khanh Tran and Duc Tinh Le, 2010):

$$A^T AZ - A^T L = 0 \quad (25)$$

Then, the result of the parameters can be solved by:

$$Z = (A^T A)^{-1} A^T L \quad (26)$$

Mono variant and multivariant regressions are applied for analyzing monitoring data of a cable - stayed bridge, including displacement data

acquired by GPS - RTK, air - temperature and wind - speed data. Variable y in each kind of regression model is chosen as the coordinate of considering point along to separate direction. Variable x depends on each case of regression: (1) in case of mono variant regression, x variable is air - temperature or wind - speed; (2) in case of multivariant regression, two variables x_1 and x_2 are denoted to air temperate and win - speed respectively. In both cases, parameter a_0 is redundant between the regression equation and the applied data. a_0 has to satisfy the condition that it is white noise with normal distribution and $\sum a_0 = 0$, its variance is constant. According to statistics, the determination parameter R^2 is used to define the appropriate regression model. It means that R^2 is closed to 1, the defined regression model is the most appropriate, the described model effectively explains the effects of variables (Sanford Weisberg, 2005; Shumway and Stoffer, 2010; Peter and Annick, 1987). Parameter R^2 is defined by:

$$R^2 = 1 - \frac{RSS}{SYY} \quad (27)$$

where RSS is the square summation of the redundant between model and data; SYY is the square summation of the deviation between the displacement i and the mean value.

This study applies a regression model for monitoring data to define parameters a_1, a_2, a_0, b , with displacement y is affected by air - temperature t and wind - speed v. Moreover, the determination coefficient R^2 is used to assess the consistency of the regression model for analyzing, assessing and predicting the specific points' displacement.

4. Experiment

4.1. Introduction of the SHM system of Can Tho cable - stayed bridge

Can Tho cable - stayed bridge was build in 2004 crossing the Hau river to connect Can Tho province to Vinh Long. Figures 2 and 3 show the target bridge and its location.

Can Tho bridge was first used in 2010 as the longest main span in the South East Asia (550 m), the total length of the main bridge is 2,750 m,

the height of the tower is 171 m. Its concrete girder is 26 m wide, but 210 middle lengths of the main span is made from steel structure.

Structural Health Monitoring System - SHMs

was established in 2010, including many kinds of sensors. It is considered a modern monitoring system in Vietnam. Figure 4 shows the sensor locations of the SHMs of Can Tho bridge.



Figure 2. Can Tho bridge location.



Figure 3. Can Tho bridge.

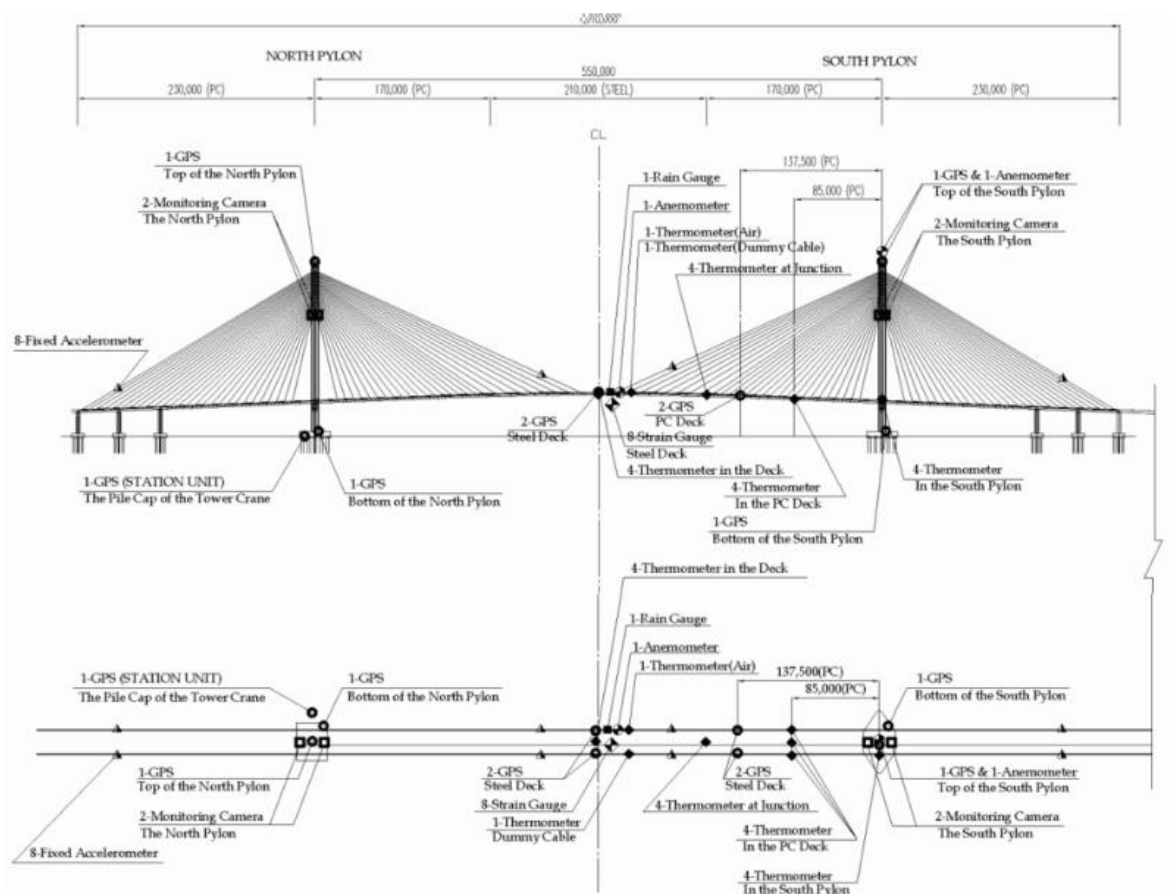


Figure 4. Diagram of sensor locations on Can Tho bridge (Farrar et al., 2000).

Global Positioning System - GPS is applied in the SHMs of Can Tho bridge, including 09 rover sensors located on some specific location such as on the top of towers, on the main girder, and other piers. Two base stations are established near the management office and near the North Pylon (Figures 5, 6, and 7). GPS equipment is Leica brand with GMX 902 version that has specific errors provided by the manufacture with ± 10 mm ± 1 ppm in the horizontal plane and ± 20 mm ± 1 ppm in the vertical direction. The principle of GPS measurement is used in SHMs is Real - Time Kinematic - RTK technique. The GPS sensor

frequency can reach 20 Hz, but the acquired GPS data are calculated to save the average value in 1 minute, 10 minute, 1 hour, and one day.

GPS technology shows various advantages in monitoring the displacements of a large - scale structure, especially in monitoring a long - span bridge. It can measure in real - time, overcoming all kinds of weather conditions, reaching to millimeter level accuracy. However, GPS technology has a bit high cost in its application, and GPS data processing is still a challenge to assess the structural health.

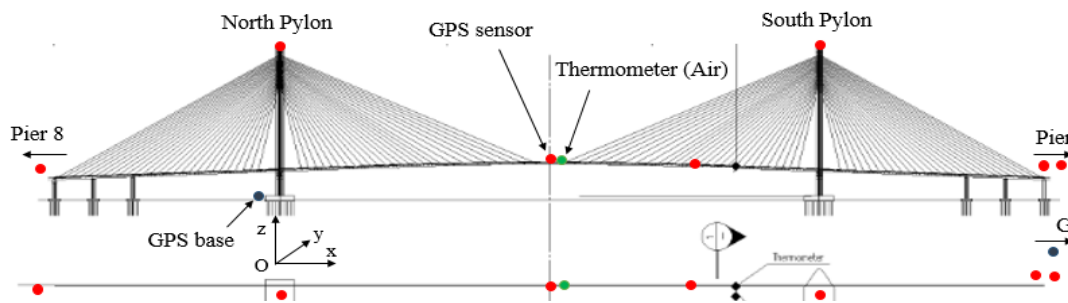


Figure 5. GPS sensors locations on Can Tho bridge.

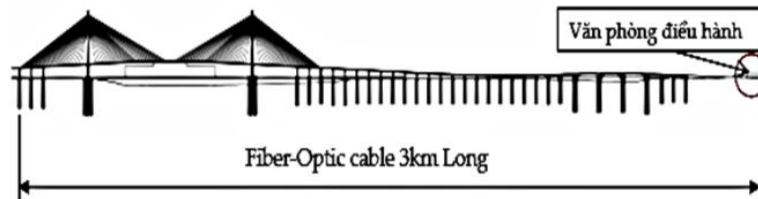


Figure 6. GPS base station location.



Figure 7. GPS rover on girder location.

4.2. Experimental data

In the monitoring system of the target bridge, all sensors acquire data in real - time at a specific frequency, then the acquired data are saved in a short time. Furthermore, the short - time data are then averaged value in 1 minute, 10 minutes, 1 hour, or 1 one day to save a long time. However, storing and analyzing long - time monitoring data is a challenge because it is a huge volume. GPS acquires the displacement monitoring data of specific points along to 3 directions: longitudinal direction (x - direction); lateral direction (y - direction), and vertical direction (z - direction).

In this experimental study, the 10 minute average data of the target bridge extracted in 3 days (from January 2nd to 5th in 2017) are used to analyze that include GPS displacement data, air - temperature data, and wind - speed data. Figure 8 shows the experimental data in 4 specific monitoring points: two points on the top of towers and two other points on the girder (at the middle of the main span and the quarter main span). Figure 9 shows the environmental data, including air - temperature and wind - speed.

4.3. Experimental Results

4.3.1. Correlation analysis

In this experimental study, the correlation between GPS displacement and environmental parameters was analyzed. Then the 95% confident interval of each correlation coefficient was also calculated. The results of correlation coefficients and the 95% confident interval of some specific direction were shown in Table 2.

The results of correlation coefficients between GPS data and environmental data show some discussion follows:

- Correlation between wind - speed and GPS data is very weak, which the correlation coefficients of all points along all directions are less than 0.5. It can be understood that the wind - speed has less effect on the displacements of the target bridge. Moreover, the wind - speed has correlated with GPS data along to the lateral direction (y - direction) that is higher than other directions (x - and z - directions). The tower points have a higher correlation with the GPS data than the girder points. This kind of result is appropriate with the characteristic of a cable - stayed bridge.

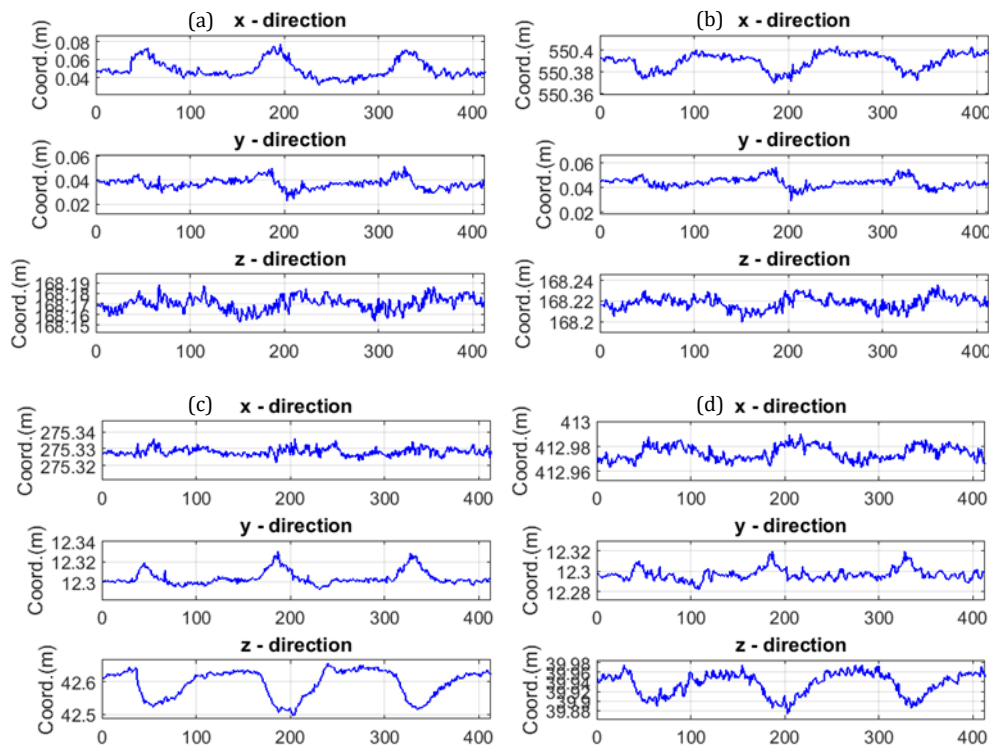


Figure 8. GPS experimental data.

(a) North tower; (b) South tower; (c) Middle span; (d) Quarter span.

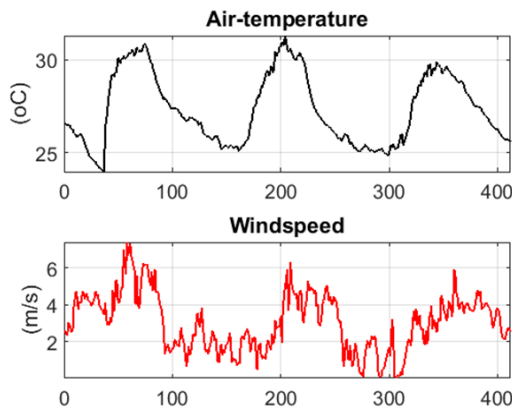


Figure 9. Environmental data.

Table 2. Results of correlation coefficients and 95% confident interval.

Points	Direction	Correlation Coefficient		95% confident interval of high coefficients
		Air - temperature	Wind - speed	
North tower points	Longitudinal - x	0.63	-0.29	0.53 ÷ 0.68
	Lateral - y	-0.35	-0.48	
	Vertical - z	0.36	0.39	
Middle span	Longitudinal - x	0.37	0.15	
	Lateral - y	0.30	-0.38	
	Vertical - z	-0.90	0.14	-0.93 ÷ -0.88
Quarter span	Longitudinal - x	0.67	0.25	0.51 ÷ 0.70
	Lateral - y	0.24	-0.24	
	Vertical - z	-0.88	0.14	-0.91 ÷ -0.86
South tower	Longitudinal - x	-0.68	0.33	-0.70 ÷ -0.58
	Lateral - y	-0.40	-0.49	
	Vertical - z	0.46	0.37	

- Correlation between the air - temperature and GPS data is very high in some specific directions, such as the vertical direction of the girder points and the longitudinal direction of the tower points. Statistically, the correlation coefficients along to the longitudinal direction (x) of the tower points are larger than 0.5, showing a reverse correlation between two points (0.63 and -0.68). The lateral and vertical directions of tower points show small correlation coefficients with the air - temperature (less than 0.5). These results show the coincidence with the movement of the bridge pylon, that they just show a significant trend along to the longitudinal direction. The correlation between air - temperature and the girder points along to vertical direction that is a very high contravariant correlation, and the coefficients are -0.90 and -0.88 for the middle and

quarter span respectively. Meanwhile, the lateral direction of the girder points shows a low correlation with the air - temperature, the longitudinal direction of the quarter span point shows a high correlation (0.67). It can be explained that the quarter span point is non - symmetric. Thus the movement of this point along to x - direction is much larger than the middle span point.

From the above discussion, the air - temperature has affected the GPS data of the target bridge. Then, the significant directions of the bridge movement can be recognized that are the longitudinal direction (x - direction) of the tower points; and the vertical direction (z - direction) of the girder points. These significant directions are then used to analyze the regression model. Moreover, the target bridge's global displacement model could be recognized through the GPS monitoring data, shown in Figure 10.

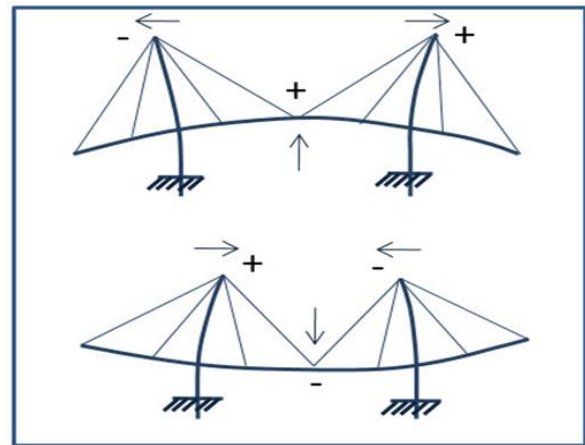


Figure 10. The global model of GPS displacement of the target bridge.

4.3.2. Regression analysis

* Establishment of mono variant regression model - Model 1

Monovariant regression model was applied for the specific points on the target bridge and the significant directions that were recognized, such as the longitudinal direction of the tower points (namely #Pt1 and #Pt4) and the vertical direction of the girder points (namely #Pt2 and #Pt3). In this application, GPS displacement data is considered a function of air - temperature variable wind - speed separately, as described in equation (28).

$$f = a_0 + a_1 t \quad (28)$$

The least - square principle was used to define the regression functions along each direction. The results of mono variant regression along with significant directions are shown in Table 3.

Figure 11 shows the mono variant regression line and the 95% confident interval of the significant directions and the determination coefficient R2 of each case. It can be seen that the mono variant regression model much coincides with the vertical direction of the girder points which the R2 coefficients are 0.80 and 0.77, respectively. Meanwhile, R2 coefficients of the tower points (#Pt1 and #Pt4) are low (0.39 and

0.47). It means that displacements of girder point (middle span and quarter span) are mainly caused by effects of temperature, and coincides with the characteristic of the target bridge.

Figure 12 shows the mono variant regression results between GPS data and wind speed data and the 95% confident interval. The R2 coefficients are also shown in the significant directions separately. It can be seen that the wind - speed has more effects on the vertical direction of the girder points than the longitudinal direction of the tower points. However, the very low R2 coefficients show that the mono variant regression with wind - speed variable is not appropriate for GPS monitoring data.

Table 3. Mono variant regression results.

Function parameters	#Pt1 (x-direction)		#Pt2 (z-direction)		#Pt3 (z-direction)		#Pt4 (x-direction)	
	Temp,	Wind-speed	Temp,	Wind-speed	Temp,	Wind-speed	Temp,	Wind-speed
a0	-0.040	0.045	43.150	42.627	40.238	39.954	550.467	550.393
a1	0.0033	0.0012	-0.0203	-0.0106	-0.0110	-0.0058	-0.0028	-0.0012

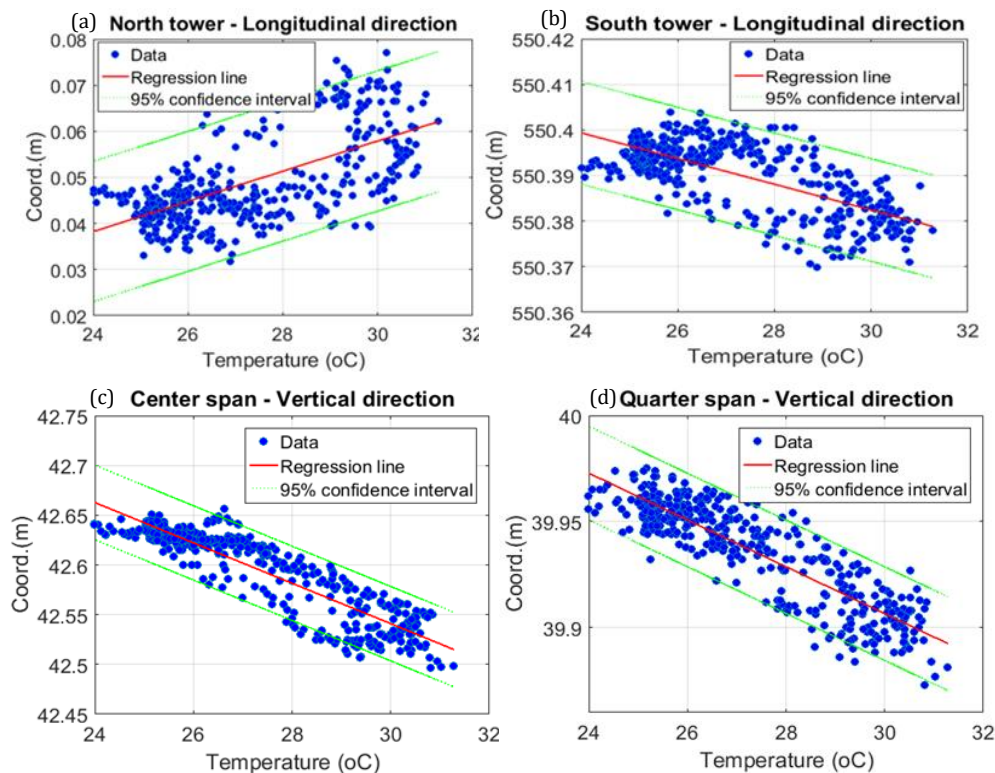


Figure 11. Monovariant regression with air-temperature variable of the significant directions.
 (a) #Pt 1 - x direction; R2 = 0.39; (b) #Pt 4 - x direction; R2 = 0.47; (c) #Pt 2 - z direction; R2 = 0.80;
 (d) #Pt 3 - z direction; R2 = 0.77;

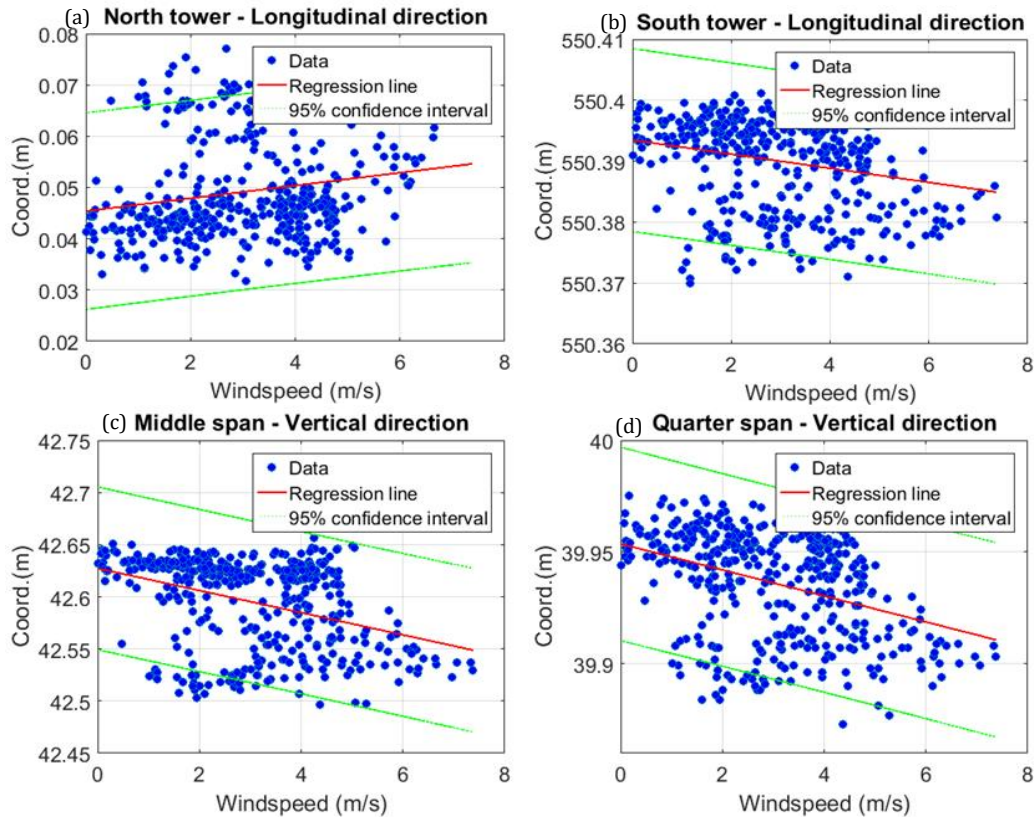


Figure 12. Monovariant regression with wind-speed variable of the significant directions. (a) #Pt 1 - x direction; $R^2 = 0.04$; (b) #Pt 4 - x direction; $R^2 = 0.05$; (c) #Pt 2 - z direction; $R^2 = 0.15$; (d) #Pt 3 - z direction; $R^2 = 0.14$.

* Establishment of multivariant regression model - Model 2

In this study, the GPS displacement along to each significant direction is considered a function of air - temperature and wind - speed variables, as described in equation (29):

$$f = a_0 + a_1 t + a_2 v \quad (29)$$

where: t is the air - temperature variable; v is the wind - speed variable.

The least - square principle was used to define the regression functions. The results of multivariant regression, along with significant directions, are shown in Table 4.

Figure 13 shows the multivariant regression plane and the 95% confident interval of the significant directions. The determination coefficient R^2 of each case is also shown in each figure. It can be seen that the wind - speed has effects on GPS monitoring data along with the significant directions of the target bridge that is

less than the air - temperature's effects, which showed in the parameters a_1 and a_2 of the regression functions (Table 4). Moreover, the air - temperature affects GPS displacement along to vertical direction of girder points approximately six times the wind - speed's effects. Similar to the mono - variant regression case, displacements of the girder points are caused mainly by the effects of temperature.

4.3.3. Statistical analysis

A regression model is considered the fitting model if it satisfies the redundant between the regression model and the real data are white noise. It means that the redundant must have the normal distribution and its p - value has to be less than 0,05, then the GPS monitoring data have statistical significance, and the regression model is appropriate to describe the displacement of a structure. Therefore, the redundant of both mono - variant and multivariant regression cases of the

experimental study was then tested in the white noise condition. The results showed that the regression models with the air - temperature variable were defined in both cases that are fitted models for GPS monitoring data. In contrast, the mono variant regression with the wind - speed variable is inappropriate for GPS displacement data of the target bridge.

5. Conclusions

The results of this study can figure out some conclusions below:

(1) GPS technology with the Real - Time Kinematic technique that can monitor the

displacements of large - scale structures, such as a long - span bridge, and GPS - RTK monitoring data can be used to assess the structural health during its operation.

(2) In this study's target bridge, GPS monitoring data have a high or very high correlation with the air - temperature monitoring data and a longitudinal direction of the tower points and vertical direction of the girder points. This studied conclusion is fitted to the target bridge's characteristic, and the global deformation of a cable - stayed bridge could be recognized based on the correlation with the air - temperature. Otherwise, the wind - speed has a

Table 4. Multivariant regression results.

Function parameters	#Pt1 (x-direction)	#Pt2 (z-direction)	#Pt3 (z-direction)	#Pt4 (x-direction)
a_0	-0.052	43.186	40.257	550.476
a_1	0.0039	-0.0220	-0.0120	-0.0033
a_2	-0.0013	0.0039	0.0021	0.0010

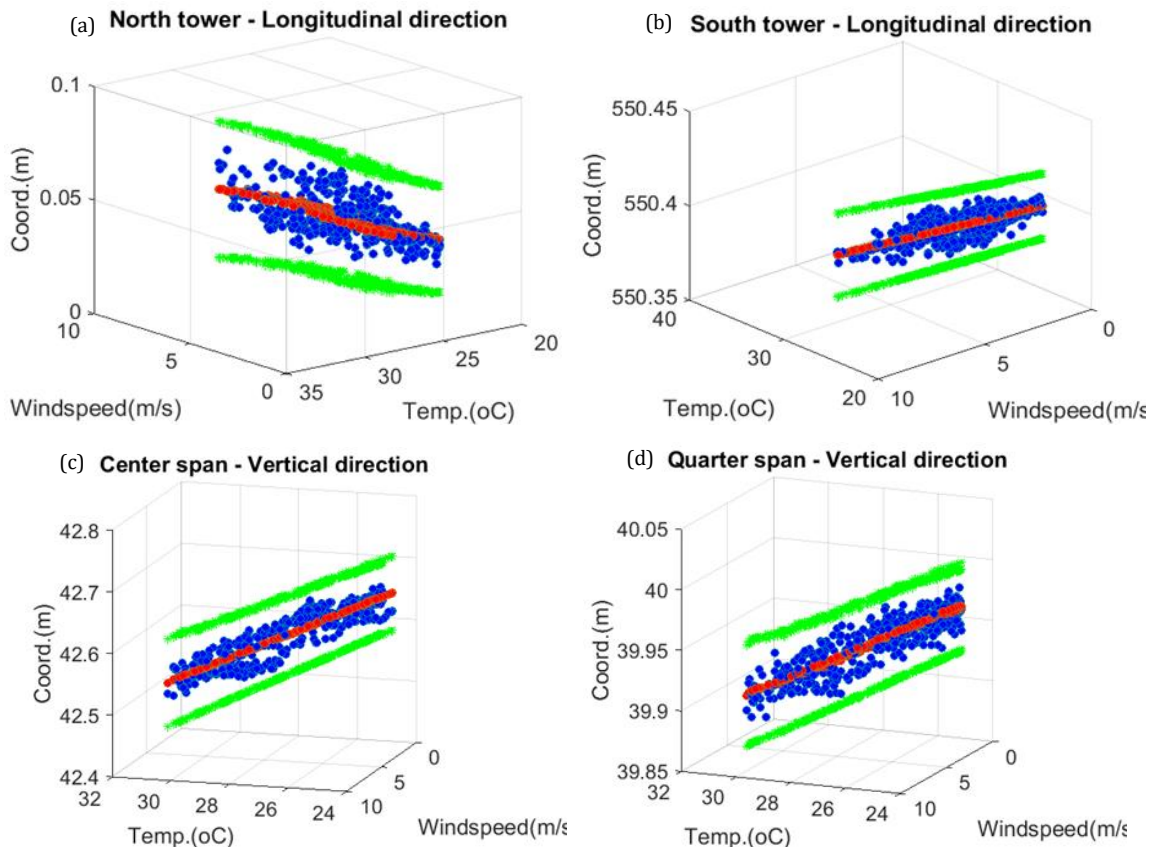


Figure 13. Multivariant regression of the significant directions

(a) #Pt 1 - x direction; $R^2 = 0.42$; (b) #Pt 4 - x direction; $R^2 = 0.49$; (c) #Pt 2 - z direction; $R^2 = 0.82$; (d) #Pt 3 - z direction; $R^2 = 0.79$.

low or no correlation with the GPS monitoring data.

(3) Regression analysis showed that the mono - variant regression with temperature variable or the multivariant regression with temperature and wind - speed variables suitable for describing the displacement model of the main span points along to vertical direction.

(4) Correlation analysis can figure out the features that affect the displacements of a long - span bridge. One is the central part causing the displacements could be recognized by the high correlation coefficients. The results of experiment shows that the correlation analysis is an effective method to analyze the GPS displacement data of the target bridges.

Overall, this study figures out how to apply the correlation and regression to assess the displacement of a cable-stayed bridge through GPS monitoring data considering the effects of air-temperature. By the proposed method, the global deformation of a cable-stayed bridge could be recognized as well as its displacements can be predicted in some significant directions under environmental temperature changes.

The author's contributions

The author Tinh Duc Le edited the abstract and results of the study; The authors Tinh Duc Le, Hien Van Le and Duy Tien Le prepared the methods and materials; The authors Thanh Kim Thi Nguyen, Linh Nguyen Thuy edited the conclusion and references.

References

Cao Van Nguyen et al., (2002). Theory of probability and mathematical statistics. *Publisher of education*, Hanoi.

Celebi, M., (2000). GPS in dynamic monitoring of long - period structures. *Soil Dynamics and Earthquake Engineering*, 20(5), 477 - 483.

Cheng, P., John, W., & Zheng, W., (2002). Large structure health dynamic monitoring using GPS technology. *In FIG XXII International Congress, Washington, DC USA*.

Cornwell, P., Farrar, C. R., Doebling, S. W., & Sohn, H., (1999). Environmental variability of modal properties. *Experimental Techniques*, 23(6), 45 - 48.

Duc Tinh Le, (2012). Application of statistical methods in analyzing the deformation of hydropower projects in Vietnam. *Final report of topic for supporting PhD student*, code number N2010 - 31.

Farrar C. R., Cornwell P., Doebling S. W., & Prime M. B., (2000). Structural Health Monitoring Studies of the Alamosa Canyon and I - 40 Bridges. *Los Alamos National Laboratory*, LA - 13635 - MS.

Fujino, Y., Murata, M., Okano, S., & Takeguchi, M., (2000). Monitoring system of the Akashi Kaikyo Bridge and displacement measurement using GPS. *In SPIE's 5th Annual International Symposium on Nondestructive Evaluation and Health Monitoring of Aging Infrastructure* (pp. 229 - 236). International Society for Optics and Photonics.

H. V. Le, M. Nishio, H. Yamada, H. Katsuchi, (2015). Statistical condition assessment of a cable stayed bridge using GPS structural health monitoring data. *The 7th International Conference of Structural Health Monitoring of Intelligent Infrastructure SHMII*, Italy.

Hien V. Le, Mayuko Nishio, (2015). Time - series analysis of GPS monitoring data from a long - span bridge considering the global deformation due to air temperature changes. *Journal of Civil Structural Health Monitoring*, Springer 5, 415 - 425.

Kalooop, M. R., & Li, H., (2009). Monitoring of bridge deformation using GPS technique. *KSCE Journal of Civil Engineering*, 13(6), 423 - 431.

Kalooop, M. R., & Li, H., (2011). Sensitivity and analysis GPS signals based bridge damage using GPS observations and wavelet transform. *Measurement*, 44(5), 927 - 937.

Khanh Tran, Tinh Duc Le, (2010). Application of correlation analysis method in assessing the structural displacement. *Journal of science and technique of mining and geology* (31).

Khanh Tran, Phuc Quang Nguyen. (2010). Structural Deformation and Displacement Monitoring. *Publisher of Transportation*.

Omenzetter P., & Brownjohn J. M. W., (2005). A seasonal ARIMAX time series model for strain - temperature relationship in an instrumented

- bridge. *Proceedings of the 5th International Workshop on Structural Health Monitoring*, 299 - 306.
- Omenzetter, P., & Brownjohn, J. M. W., (2006). Application of time series analysis for bridge monitoring. *Smart Materials and Structures*, 15(1), 129 - 138.
- Peter J. Rousseeuw, Annick M. Leroy, (1987). Robust Regression and Outlier Detection. Book, Wiley - Interscience.
- Sanford Weisberg, (2005). Applied Linear Regression. Book, third edition. Wiley - Interscience.
- Shumway, R. H., Stoffer, D. S., (2010). Time series analysis and its applications: with R examples. Springer.
- Sohn, H., Czarnecki, J. A., Farrar, C. R., (2000). Structural health monitoring using statistical process control. *Journal of Structural Engineering*, 126(11), 1356 - 1363.
- Sohn, H., Dzwonczyk, M., Straser, E. G., Kiremidjian, A. S., Law, K. H., & Meng, T., (1999). An experimental study of temperature effect on modal parameters of the Alamosa Canyon Bridge. *Earthquake engineering & structural dynamics*, 28(8), 879 - 897.