International Journal of Engineering Innovation & Research
Volume 8, Issue 3, ISSN: 2277 – 5668

The Application of Time Series Analysis in Forecasting Bridge Displacement Data: Case Study in Wuxi Bridge, Taiwan

Hieu Ha Trung[1*], Tien Yin Chou[2] and Yao Min Fang[2]
[1*]Ph.D program of Civil and Hydraulic Engineering, College of Construction and Development, Feng Chia University, Taiwan.
[2]GIS Research Center, Feng Chia University, Taichung, Taiwan.
*Corresponding author email id: hatrunghieu0409@gmail.com

Date of publication (dd/mm/yyyy): 03/07/2019

Abstract **– Time series analysis is a term referring to the process of fitting a time series dataset to a suitable model. An Auto-Regressive Integrated Moving Average model is introduced in this paper and then applied to a case study, which is forecasting bridge monitoring displacement data. To evaluate the forecast performances, root mean square errors and mean absolute percentage error are utilized. The results shown that ARIMA (2, 1, 2) model can be a potential and effective model in predicting the daily future data in such case study.**

Keywords **– Time Series Analysis, Bridge Displacement, Time Series Data, Wuxi Bridge, Forecasting Data.**

I. INTRODUCTION

Time series analysis is a term that refers to the procedure of fitting a time series data to a suitable model. This model is then commonly used to forecast future values for the series by understanding the historical observations [1] [2]. Numerous time series models have been proposed over the last decades, in which these models can be linear or non-linear: Linear models such as Auto-Regressive (AR) models, Moving Average (MA) models, Auto-Regressive Moving Average (ARMA), Auto-Regressive Moving Average with exogenous input (ARMX), Auto-Regressive Integrated Moving Average (ARIMA). Non-linear models such as artificial neural network, Support Vector Machine (SVM), generalized regression neural networks (GRNNs) and multi-layer feed forward networks (MLFNs) are widely popular used [3] [4] [5].

One of the most frequently used time series models is the ARIMA model [6] [7]. It can be applied in many fields such as business, economics, industry, engineering and science due to the popularity of time series data in nature. Recently, an ARIMA model has developed to forecast value of rice price fluctuations in Indonesia [8]. Lanyi Zhanget al. applied an ARIMA model for forecasting Fine airborne particles ($PM_{2.5}$) concentrations to have a better understanding in air quality improvement in China [9]. In future sales performances, the prediction of consumer up-coming data is demonstrated through a case study of retail sales of women footwear by comparing state space models and ARIMA models [10]. The combination of GIS, spatial autocorrelation index and ARIMA model to estimate the channel shifting in Bangladesh in the period of 1972–2031 has been proposed [11]. Arumugamet alused rainfall time series data in 10 years from 2006 to 2016 to build an ARIMA model for forecasting future precipitation [12]. A case study in South Korea was applied seasonal autoregressive integrated moving average model to predict future solar radiation based on the hourly radiation observations in almost 40 years [13].

Displacement is one of the important factors presenting bridge health assessment. To measure displacement of a bridge, the most popular method used is install sensors on different reference positions but it is difficult to deploy the sensors at some fixed reference points on bridge [14]. Several studies have been published presenting different methods for estimating bridge displacement [15-19]. During operation, displacement of bridge always has to be observed and gradually checked due to the direct relation to the structural stiffness and its integrity. In some cases, bridge displacement values are missed because of equipment errors or the impact of weather conditions thus it is necessary to correctly predict the missing data or forecast the future observations. This paper introduces a case study with a proper model for forecasting displacement data. Root mean square errors (RMSE) and mean absolute percentage error (MAPE) are utilized to judge forecasting performances of different methods.

II. Theoretical Background

2.1. *ARIMA Model*

An ARIMA model was first introduced by Box and Jenkins [6]. The general equation of successive differences at the d^{th} difference of X_t is as follows [9] [13] :

$$\Delta^d X_t = (1 - B)^d X_t \tag{1}$$

Where d refers to the difference order where and B denotes the backshift operator.

The successive difference at one-time lag is computed as:

$$\Delta^1 X_t = (1 - B) X_t = X_t - X_{t-1} \tag{2}$$

In this case study, the general ARIMA (*p,d,q*) is expressed as:

$$\Phi_p(B) W_t = \theta_q(B) e_t \tag{3}$$

Where $\Phi_p(B)$ is an autogressive operator of order p, $\theta_q(B)$ is a moving average operator of order q, and $W_t = \Delta d X_t$

Generally, an ARIMA (p, d, q) model contains three terms that are autoregressive term, integrated term, and moving average term with the p, d, and q orders respectively. To develop an ARIMA model, the first step is to check whether the time series data is stationary or non-stationary, which can be diagnosed by Augmented Dickey-Fuller (ADF) test. If there is a trend or seasonality in time series data, it must be made stationary by differencing. Next step, use the Autocorrelation function (ACF) and partial autocorrelation function (PACF) to decide whether to include an AR term (s), MA term (s), or both.

The function of ACF and PACF can be respectively illustrated as follow [9]:

$$R(s,t) = \frac{E[(X_a - \mu_a)(X_b - \mu_b)]}{\sigma_a \sigma_b} \tag{4}$$

$$\begin{aligned} \alpha(1) &= cor(X_{t+1}, X_t), k = 1 \\ \alpha(k) &= cor(X_{t+k} - P_{t,k}(X_{t+k}), X_t - P_{t,k}(X_t)), k \geq 2 \end{aligned} \tag{5}$$

ARIMA model in this paper was developed using Matlab software; this software was also utilized to calculate the ACF and PACF.

2.2. *Augmented Dickey-Fuller Test*

An Augmented Dickey-Fuller test (ADF test) of a series is carried out under the null hypothesis $H_0 = 0$ against the alternative hypothesis $H_1 < 1$. A value for test statistic is calculated as:

$$DF_i = \frac{\hat{\gamma}}{SE(\hat{\gamma})} \tag{6}$$

Where: $\hat{\gamma}$ indicates the least square estimate; $SE(\hat{\gamma})$ denotes the usual standard error estimate. The test statistic value then compared with the critical values for the Dickey-Fuller test and be decided if the null hypothesis is accepted or rejected.

2.3. *Root mean Square Errors and mean Absolute Percentage Error*

The best fitting model to the displacement data in this paper is evaluated by root mean square errors (RMSE) and mean absolute percentage error (MAPE) [20] [21] [22]:

$$RMSE = \sqrt{\sum_{1}^{n} \frac{(X_t - X_0)^2}{n}} \tag{7}$$

$$MAPE = \frac{1}{n} \sum_{1}^{n} \left| \frac{(X_t - X_0)}{X_0} \right| \times 100 \tag{8}$$

Where X_t, X_0 are the forecasted data and observed data, respectively; n is the total observations.

III. The Application of ARIMA Model in Forecasting Bridge Displacement Data

3.1. *Displacement Time Series Data in Wuxi Bridge*

Wuxi bridge line 63 (Fig. 1) is an important concrete bridge connecting Taichung City and Nantou County in Taiwan. Due to the increase in extreme climate events because of global climate change, the foundational erosion of the bridge has become more serious. Therefore, a monitoring system has been installed on this bridge with the purposes of monitoring bridge conditions: Measuring and evaluating the load effect to the bridge, using the actual measurement of the vibration characteristics to evaluate structural performance; evaluating the environment impact to the bridge and maintenance management. Figure 2 and Figure 3 respectively show the cross-section of Wuxi bridge and the settlements gauges deployment in one cross-section.

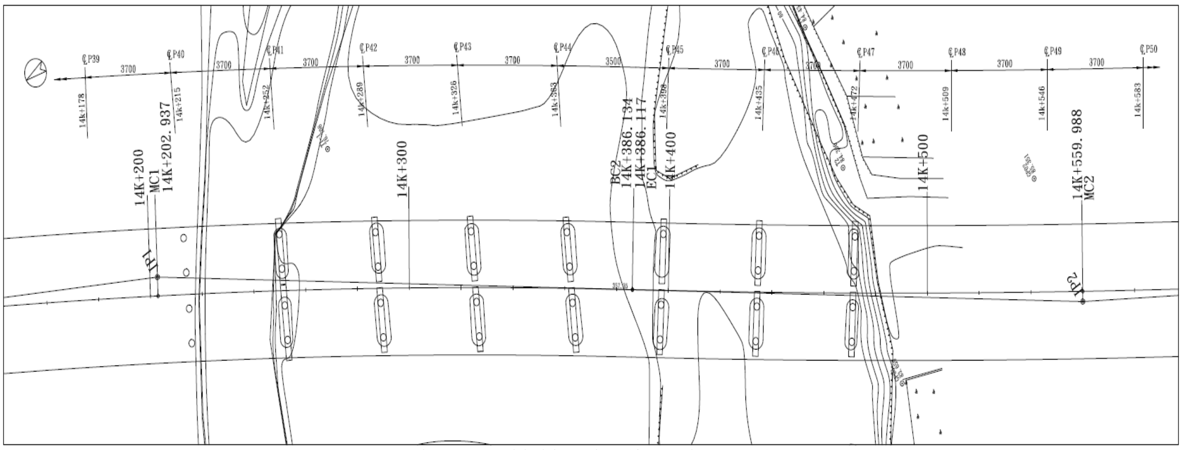


Fig. 1. Wuxi bridge plan (from pier 42-48)

Fig. 2. Cross-section of bridge

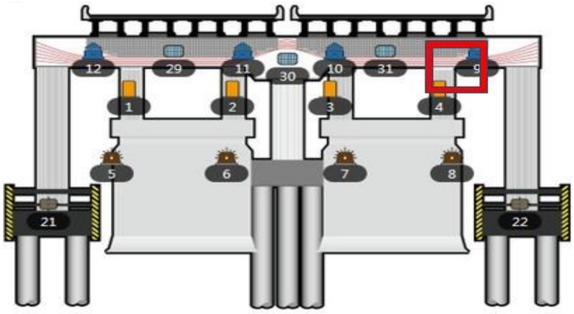


Fig. 3. Settlement gauges in one section

The data used in this paper is ten-minute average displacement time-series data, which extracted in 90 days (from March 8th 2018 to June 4th 2018), the ten-minute data then be averaged to daily displacement observations. Figure 4 show the plots of acquired displacement time-series data at the Pier 47 and the daily displacement data is plotted in Figure 5.

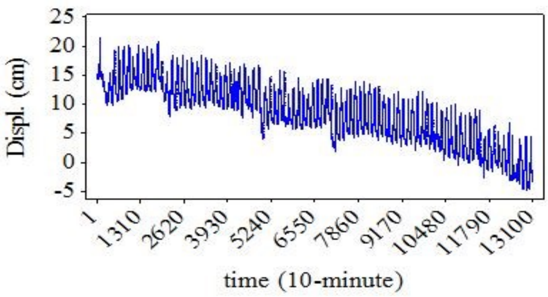


Fig. 4. 10-minute displacement data

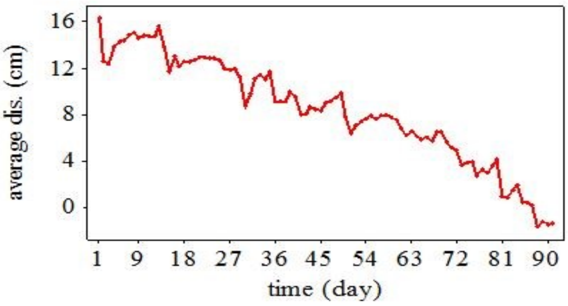


Fig. 5. Daily displacement data

3.2. *Application of ARIMA Model in Forecasting Displacement Time Series Data*

The daily dataset has total 90 observations. Out of these the first 75 (data in 75 days) are considered for training and the remaining 15 for testing.

The ADF test is applied to check the stationarity of dataset. In the theory of ADF test, the statistical possibility (p-value) and the absolute critical value are the key to decide the null hypothesis of a test is accepted or rejected. The null hypothesis of ADF test here is assumed as "the dataset is non-stationary". The result of this test shows that, the possibility value (p-value) is 0.1361 (13.61 %) that is greater than the 5% of significant level. Therefore, the null hypothesis here is accepted, it indicates that the dataset timeseries is non-stationary. In order to make dataset stationary, the first order of difference has been taken and ADF test has been rerun. The result shows that the p-value is 0.001 (0.1%) thus the time series data is considered as stationary. Figure 6 below shows the daily displacement data after differencing.

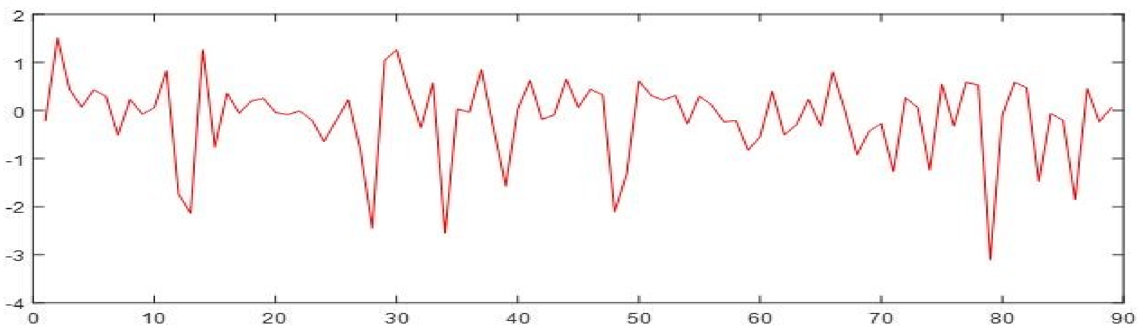
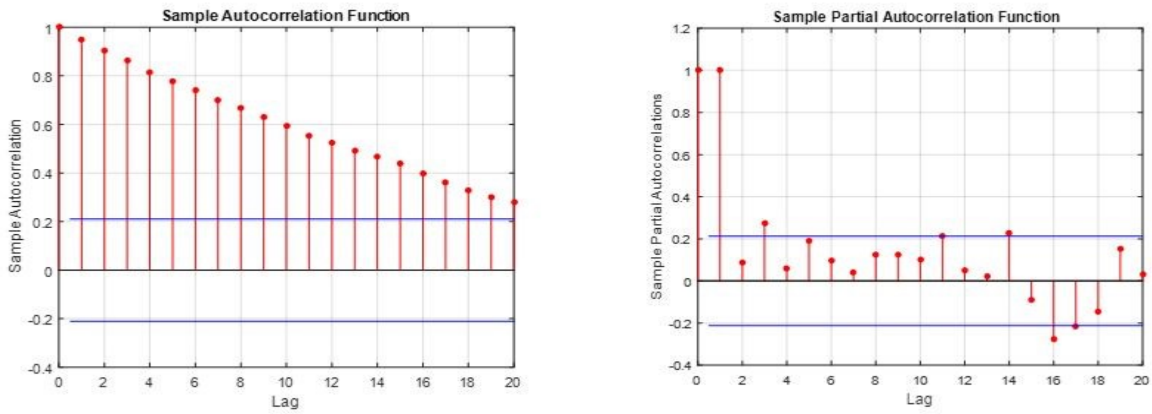


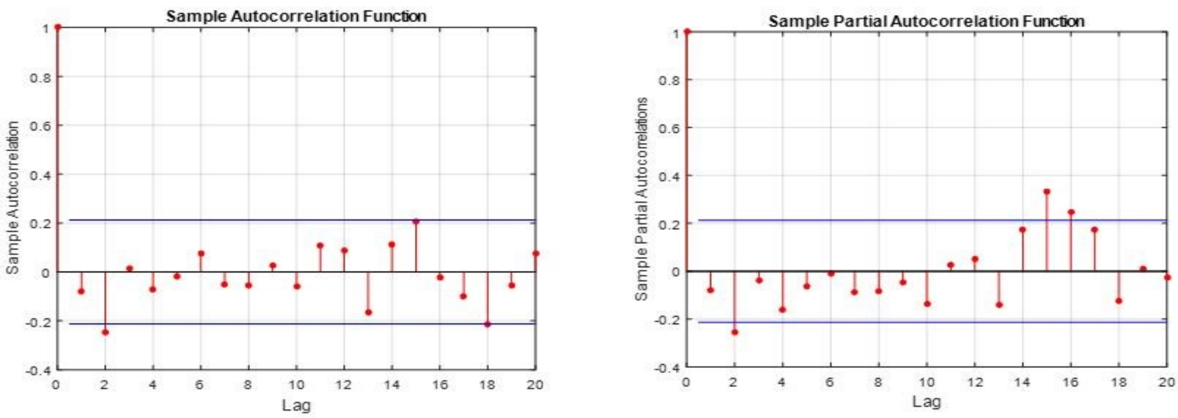
Fig. 1. Differentiated data

The ACF and PACF are calculated by *autocorr* and *parcorr* function in Matlab. ACF and PACF of dataset before and after differencing are shown in Figure 7 and Figure 8 as follow:



a) ACF b) PACF

Fig. 2. a, b. ACF and PACF of displacement time series data



a) ACF b) PACF

Fig. 8. a, b. ACF and PACF of differentiated data

Figure 8 shows that almost the values from lag 2 in both ACF and PACF of differentiated data are inside the 95% confidential internal which is marked by blue lines. Thus the fitted model are assumed to ARIMA (2, 1, 2).

It also can be intuitively seen that the daily displacement dataset is a linear approximation. Thus, the linear model and quadratic model are also considered to make comparisons with ARIMA model. These models are estimated by trend analysis in Minitab software.

The three forecasting plots are shown in Figures below:

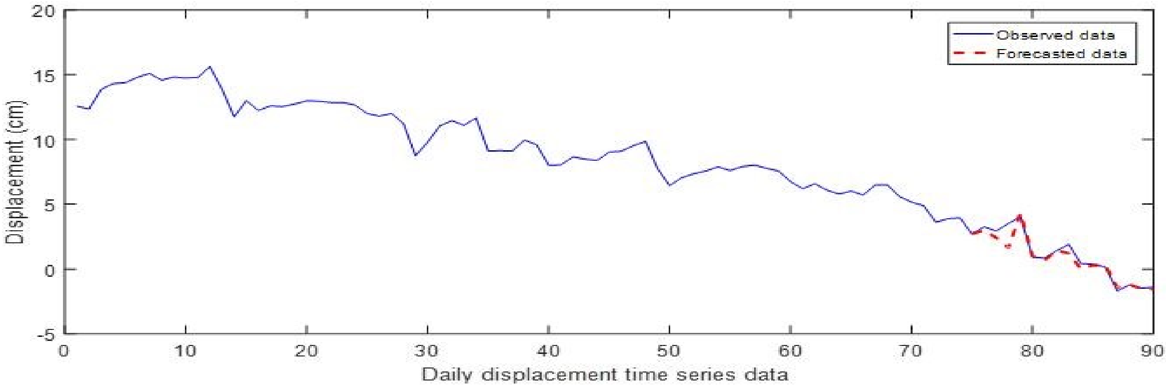


Fig. 9. ARIMA (2, 1, 2) forecast diagram

International Journal of Engineering Innovation & Research
Volume 8, Issue 3, ISSN: 2277 – 5668

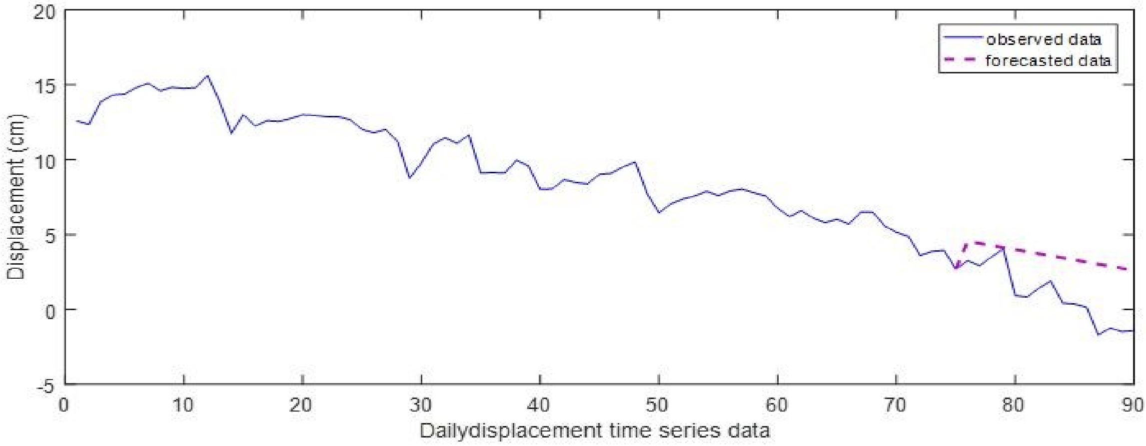


Fig. 10. Linear model forecast diagram

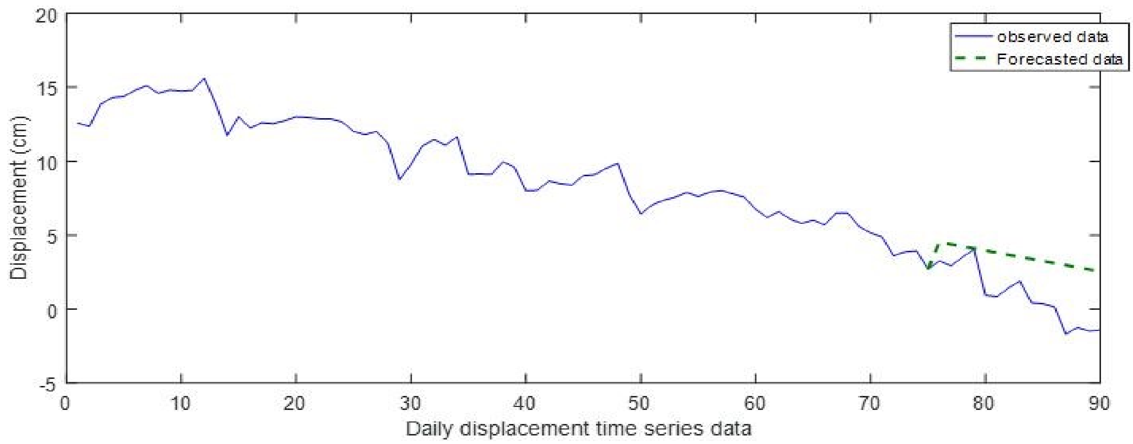


Fig. 11. Quadratic model forecast diagram

The performance of forecasting displacement time series data are computed and shown in Table 1:

Table 1. Forecast Results for Daily Displacement Time Series

Method	RMSE	MAPE (%)
ARIMA (2, 1, 2)	0.231	4.45
Linear model ($Y_t = 15.872 - 0.167037t$)	1.207	62.84
Quadratic model ($Y_t = 14.339 - 0.0671t0.00109t^2$)	1.188	61.79

It can be seen from the above table that the best fitted model is obtained by using ARIMA (2, 1, 2) model with the smallest errors in forecasting.

IV. Conclusion

Forecasting in time series data is an attractive research field as well as such a promising future works. A case study of forecasting displacement time series observations is introduced in this paper. Some statistical tools such as Augmented Dickey-Fuller test, Root mean square errors (RMSE) and mean absolute percentage error (MAPE) are also presented. The results has shown that the ARIMA (2, 1, 2) model is the best fitting model to the daily displacement time series data with the values of RMSE and MAPE are equal to 0.231 and 4.45(%), respectively. Therefore it can be a potential and effective model in predicting the daily future data in such case study.

REFERENCES

[1] Peter J. Brockwell Richard A. Davis, *Introduction to Time Series and Forecasting - Second Edition*, Springer texts in statistics (2002), ISBN 0-387-95351-5.

[2] David R. Brillingern, *Time Series: General*, Int. Encyc. Social and Behavioral Sciences (2000).

[3] Ina Khandelwal, Ratnadip Adhikari, Ghanshyam Verma, *Time Series Forecasting using Hybrid ARIMA and ANN Models based on DWT Decomposition*, International Conference on Intelligent Computing 2015, Communication & Convergence.

[4] S.N. Singh, Abheejeet Mohapatra, *Repeated wavelet transform based ARIMA model for very short-term wind speed forecasting*, Renewable Energy 136 (2019) 758e768, 0960-1481.

[5] María Victoria Riesgo García et al, *Rare earth elements price forecasting by means of transgenic time series developed with ARIMA models*, Resources Policy 59 (2018) 95–102, 0301-4207.

[6] G.E.P. Box, G. Jenkins, *Time Series Analysis, Forecasting and Control*, 1970, Holden-Day, San Francisco, CA.

[7] S.L. Ho and M. XIE, *the use of ARIMA models for reliability forecasting and analysis*, Computers ind. Engng Vol. 35 (1998), Nos 1-2, pp. 213-216.

[8] Margaretha Ohyvera, Herena Pudjihastutib, *Arima Model for Forecasting the Price of Medium Quality Rice to Anticipate Price Fluctuations*, 3rd International Conference on Computer Science and Computational Intelligence 2018, 1877-0509.

[9] Lanyi Zhang, et al, *Trend analysis and forecast of PM2.5 in Fuzhou, China using the ARIMA model*, Ecological Indicators 95 (2018) 702–710, 1470-160X/ © 2018 Elsevier Ltd.

[10] Patrícia Ramos, Nicolau Santos, RuiRebelo*, Performance of state space and ARIMA models for consumer retailsales forecasting*, Robotics and Computer-Integrated Manufacturing (2015), https://doi.org/10.1016/j.rcim.2014.12.015.

[11] Shumona Akhteret *al*, *Predicting spatiotemporal changes of channel morphology in the reach of Teesta River, Bangladesh using GIS and ARIMA modeling*, Quaternary International Volume 513, 20 April 2019, Pages 80-94.

[12] P. Arumugam and R. Saranya, *Outlier Detection and Missing Value in Seasonal ARIMA Model using Rainfall Data*, Materials Today: Proceedings Volume 5, Issue 1, Part 1, 2018, Pages 1791-1799.

[13] Mohammed H. Alsharif, Mohammad K. Younes and Jeong Kim, *Time Series ARIMA Model for Prediction of Daily and Monthly Average Global Solar Radiation: The Case Study of Seoul, South Korea*, Symmetry 2019, 11, 240.

[14] Jun-Chang Jeon, Hee-Hyun Lee, *development of displacement estimation method of girder bridges using measured strain signal induced by vehicular loads*, Engineering Structures 186 (2019) 203–215.

[15] David Hester et al, *Low- cost bridge load test: Calculating bridge displacement from acceleration for load assessment calculations*, Engineering Structures 143 (2017) 358–374.

[16] K.-T. Park et al., T*he determination of bridge displacement using measured acceleration*, Engineering Structures 27 (2005) 371–378.

[17] S.M. Seyed Ardakani, M.S. Saiidi, *Simple method to estimate residual displacement in concrete bridge columns under near-fault earthquake motions*, Engineering Structures 176 (2018) 208–219.

[18] J. Kim et al), *Autonomous dynamic displacement estimation from data fusion of acceleration and intermittent displacement measurements*, Mechanical Systems and Signal Processing 42 (2014) 194–205.

[19] A.I. Ozdagli et al., *Low-cost, efficient wireless intelligent sensors (LEWIS) measuring real-time reference-free dynamic displacements*, Mechanical Systems and Signal Processing 107 (2018) 343–356.

[20] Davor Z. Antanasijević et al., *PM_{10} emission forecasting using artificial neural networks and genetic algorithm input variable optimization*, Science of the Total Environment 443 (2013) 511–519.

[21] Paul M. Swamidass, *Encyclopedia of Production and Manufacturing Management*, (2000) 978-1-4020-0612-8.

[22] Coskun Hamzaçebi, *Improving artificial neural networks' performance in seasonal time series forecasting*, Information Sciences 178 (2008) 4550–4559.

AUTHOR'S PROFILE

HieuHa Trung
PhD student at Ph.D program of Civil and Hydraulic Engineering, College of Construction and Development, Feng Chia University, Taiwan. Research interests: Civil Engineering, bridge monitoring, Geoinformatics