ARIMA MODEL TIME-SERIES FORECASTING FOR STRUCTURAL MONITORING USING RTK-GPS

Siti Aminah Anshah, Anuar Ahmad & Zulkarnaini Mat Amin

Department of Geomatic Engineering,

Faculty of Geoinformation Science. & Engineering

Universiti Teknologi Malaysia

81 310 UTM Skudai, Johor Bahru, MALAYSIA

Email: mymeen_83@yahoo.com & anuarahmad@utm.my

Abstract

Many studies have been reported by researchers on the deployment of high precision GPS sensors on large engineering structures such as dams, bridges, towers and tall building to provide real time measurement for the indication of displacements and vibrations caused due to temperature changes, wind loading, distant earthquakes, landslides, etc. Similarly, current researches on Global Positioning System (GPS) and its applications to structural monitoring have been conducted but eventually no detailed or thorough studies on the analysis of the various changes in unusual events as well as structure change or damage have been discussed or explained. As consequence, a new analysis technique has been proposed in this paper. This technique analyzed the response of the object or structure's response in the time domain. In this paper, a time series algorithm is presented for damage identification and forecasting to detect any movement of the structure. The vibration signals obtained from GPS are modelled as autoregressive integrated moving average (ARIMA) time series. The Box-Jenkins methodology of forecasting was used and it is different from the other methods because it does not assume any particular pattern in the historical data of the series to be forecasted. It uses an iterative approach of identifying a possible model from a general class of models. The chosen model is then checked against the historical data to see whether it accurately describes the series. The model fits well if the residuals are generally small, randomly distributed, and contain no useful information. By using Minitab software, this Box-Jenkins methodology can be implemented in the model building strategy and the model can be used for forecasting.

1. Introduction

In the last few years, structural health monitoring has become of great significance in the civil engineering community Chang (2001). There is need to continuously monitor the level of performance and safety of structures, while subjected to everyday loads as well as earthquakes, hurricanes and other extreme events, due to increased safety requirements and financial implications. Recent research has demonstrated that wireless sensing networks can be successfully used for structural monitoring (Straser & Kiremidjian, 1998) and (Lynch, *et.a*l; 2004). Most currently available damage detection methods are global in nature, i.e., the dynamic properties (natural frequencies and mode shapes) are obtained for the entire structure from the input-output data using a global structural analysis have been described by Doebling, *et al.*, 1996.

In the past few years, methods have been developed that utilise statistical signal processing techniques to identify damage such in Sohn, *et al*., 2001 and Lei, *et. al*; 2003. Such methods rely on signatures obtained from the recorded vibration, strain or other data to extract features that change with onset of damage. These features can then be classified in a pattern classification framework. A pattern classification algorithm involves the following steps: (i) the evaluation of the structure's operational environment, (ii) the acquisition of the structural response measurements and data processing, (iii) the extraction of features that are sensitive to damage, and (iv) the development of statistical models for feature discrimination. These methods avoid the complexity of global system identification techniques and are particularly suitable for ongoing monitoring purposes. In this study, simulation of shaker-bridge was use to obtained the data using method of Real Time Kinematic Global Positioning System (RTK-GPS). In this simulation started with the antenna static for a few seconds to get undamaged case. After data of undamaged case was recorded then introduced vibration to the bridge to get the damaged case of data and record that data again.

The paper summarizes the autoregressive integrated moving average (ARIMA) modelling aspects of the vibration signals. Then implement the model building strategy by Box-Jenkins methodology. In this methodology contain model identification, model estimation and model checking that are calculating automatic from the Minitab program. Some preliminary results are presented.

2.**0 Description of the Damage Algorithm**

Structural damage affects the dynamic properties of a structure, resulting in a change in the statistical characteristic of the measured acceleration time histories. Thus damage detection can be performed using time series analysis of vibration signal measured from a structure before and after damage. In this study we use ARIMA time series to model the vibration data obtained from GPS sensor. Model ARIMA were use to forecast the future values of an observed time series of GPS data. If the GPS data exceeded from the limits of range of the ARIMA model (upper limit and lower limit) so further inspections need to be done to check the movements of the structure.

2.1 Modelling of the Vibration Signals

A typical vibration signal from GPS sensor is shown in Figure 1. Before fitting the ARMA model to the GPS data, it is important to remove the trend in order to make sure the data is stationary. The first step is to determine whether the series is stationary that is whether the time series appears to vary about fixed level. It is useful to look at a plot of the series along with the sample autocorrelation function. Easily, a non-stationary time series is indicated if the series appears to grow or decline over the time and the sample autocorrelation fail to die out rapidly. Model of stationary series are called autoregressive integrated moving average and denoted by ARIMA(p,d,q) which is p indicates the order of the autoregressive (AR) part, d indicates the amount of differencing and q indicates the order of the moving average (MA) part. Equation (1) referred to autoregressive part while equation (2) referred to moving average part. Both of these equations can combine to be auto regressive moving average equation. The p^{th}-order autoregressive takes the form;

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_2 Y_{t-p} + \varepsilon_t \quad (1)$$

where,

Y_t = response (dependent) variable at time t

Y_{t-1}, Y_{t-2}, Y_{t-p} = response variable at time lags t-1, t-2, t-p, respectively: these Y's play the role of independent variables

□$_0$, □$_1$,…□$_p$ = coefficients to be estimated

ε_t = error term at time t that represents the effects of the variables not explained by the model

while the q^{th}-order moving average model takes the form;

$$Y_t = \mu + \varepsilon_t - \omega_1 \varepsilon_{t-1} - \omega_2 \varepsilon_{t-2} - \ldots - \omega_q \varepsilon_{t-q} \quad (2)$$

where,

Y_t = response (dependent) variable at time t

μ = constant mean of the process

$\omega_1, \omega_2, \ldots \omega_q$ = coefficients to be estimated

ε_t = error term at time t that represents the effects of the variables not explained by the model

$\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots \varepsilon_{t-q}$ = errors in previous time periods that, at time t, are incorporated in the response Y_t

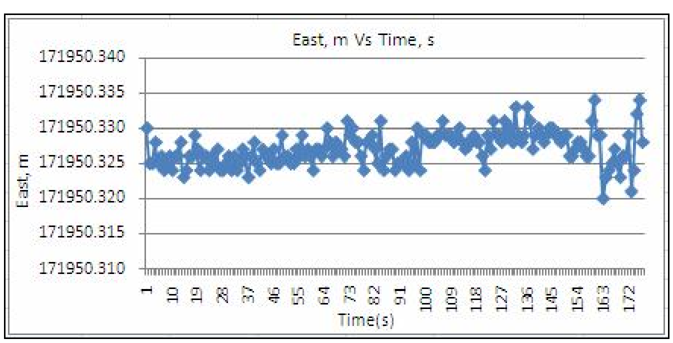

Figure 1**:** A typical raw signal time history from undamaged case

Once a stationary series has been obtained then identify the form of the model to be used. It can be done with comparing the autocorrelation function (acf) and partial autocorrelation function (pacf) computed from the data to the theoretical Hanke and Wichern, 2005) for help in selecting appropriate model. Each ARIMA model has a unique set of acf and pacf relations so to select an adequate model; Table 1 is the summary of acf and pacf patterns for autoregressive-moving average processes. Next, the order of the MA and AR parts by counting the number of significant acf and pacf, while to judge their significance both acf and pacf are usually compared with $\pm 2/\sqrt{n}$ where n is the number of observations in the time series.

Table 1: Summary of acf and pacf patterns for autoregressive-moving average processes

(Hanke and Wichern, 2005)

	Autocorrelations	**Partial Autocorrelations**
MA(*q*)	Cut off after the order *q* of the process	Die out
AR(*p*)	Die out	Cut off after the order *p* of the process
ARMA(*p*,*q*)	Die out	Die out

According to Box-Jenkins procedures (Box, *et al*.; 2000), parameters in ARIMA models are estimated by minimizing the sum of squares of fitting errors which is can be obtained using nonlinear least square procedure. A nonlinear least squares procedure is simply an algorithm that finds the minimum of the sum of squared errors function. Values of *t* can be constructed and interpreted in the usual way after estimate least squares and determined standard errors. Parameters that are judged significantly differently from zero are retained in the fitted model; parameters that are not significant are dropped from the model.

3. Application ARIMA Model for Forecasting to Detect Structural Movement

3.1 Analysis Signal for Forecasting

This studied began with collected data from the sensor of GPS. The readings after remove mean are shown in Table 2 and the analysis began by looking at a plot of the data shown in Figure 2. There appeared to be up and down trend in the series and the next step in identifying a tentative model was to look at the sample autocorrelation of the data shown in Figure 3. The autocorrelation shows that the several autocorrelation were persistently large and trailed off to zero rather slowly and the observation was correct and that this time series was non-stationary and did not vary about a fixed level. This data need to different to eliminate the trend and create a stationary series. A plot of the differenced data shown in Figure 4 and it appears to vary about a fixed level. To form the time series model it can refer from acf (shown in Figure 4) and pacf (shown in Figure 5) following to the Table 1. Next, compare the acf and pacf to the theoretical as describes before this.

Table 2: Readings for GPS data

t(s)	E(m)	t(s)	E(m)	t(s)	E(m)	t(s)	E(m)	t(s)	E(m)	t(s)	E(m)	t(s)	E(m)	t(s)	E(m)	t(s)	E(m)	t(s)	E(m)
447	0.330	463	0.326	479	0.324	495	0.329	511	0.330	527	0.329	543	0.330	559	0.328	575	0.329	591	0.330
48	0.325	464	0.329	480	0.325	496	0.326	512	0.328	528	0.327	544	0.324	560	0.327	576	0.330	592	0.329
449	0.325	465	0.327	481	0.327	497	0.326	513	0.326	529	0.325	545	0.329	561	0.328	577	0.328	593	0.329
450	0.328	466	0.324	482	0.326	498	0.325	514	0.328	530	0.331	546	0.329	562	0.328	578	0.333	594	0.328
451	0.325	467	0.326	483	0.323	499	0.325	515	0.327	531	0.324	547	0.328	563	0.329	579	0.329	595	0.329
452	0.326	468	0.326	484	0.326	500	0.327	516	0.327	532	0.326	548	0.328	564	0.328	580	0.328	596	0.329
453	0.324	469	0.324	485	0.328	501	0.326	517	0.326	533	0.327	549	0.328	565	0.328	581	0.329	597	0.326
454	0.325	470	0.326	486	0.325	502	0.329	518	0.331	534	0.327	550	0.329	566	0.326	582	0.333	598	0.326
455	0.326	471	0.325	487	0.324	503	0.326	519	0.329	535	0.324	551	0.329	567	0.324	583	0.331	599	0.327
456	0.324	472	0.327	488	0.327	504	0.327	520	0.330	536	0.325	552	0.331	568	0.329	584	0.327	600	0.328
457	0.326	473	0.324	489	0.326	505	0.326	521	0.328	537	0.325	553	0.329	569	0.327	585	0.329		
458	0.326	474	0.324	490	0.326	506	0.324	522	0.328	538	0.325	554	0.329	570	0.331	586	0.330		
459	0.328	475	0.325	491	0.325	507	0.327	523	0.326	539	0.326	555	0.329	571	0.329	587	0.329		
460	0.323	476	0.326	492	0.327	508	0.327	524	0.324	540	0.324	556	0.328	572	0.329	588	0.328		
461	0.324	477	0.324	493	0.325	509	0.326	525	0.328	541	0.328	557	0.329	573	0.328	589	0.329		
462	0.326	478	0.326	494	0.325	510	0.327	526	0.328	542	0.325	558	0.330	574	0.331	590	0.330		

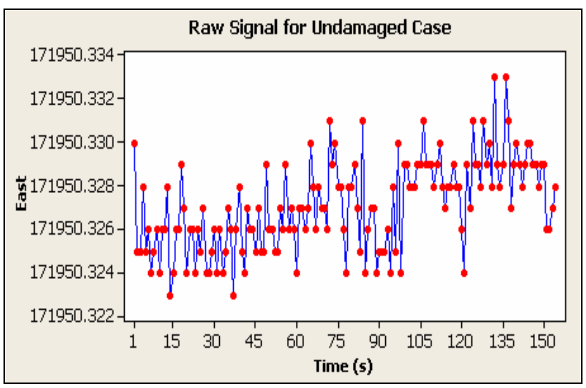

Figure 2: Raw Signals from Undamaged Case

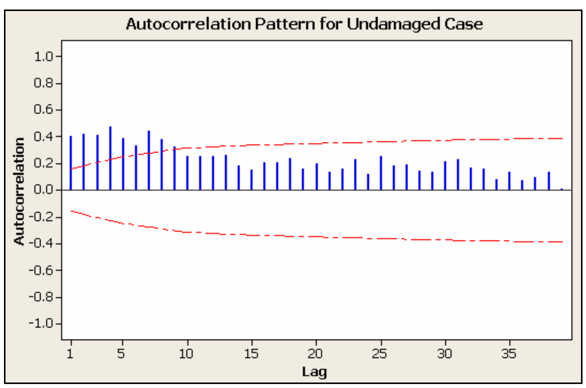

Figure 3: Sample Autocorrelation for Undamaged Case

Again the order of the MA and AR parts can determined by comparing the autocorrelation and partial autocorrelation with their error limits which is denoted by red colour dotted lines (refer to Figure 4 and Figure 5). The only significant autocorrelation was at lag 1 while for partial autocorrelation the significant at lag 3. The autocorrelations appear to cut off after lag 1, indicating MA(1) behaviour. At the same time, the partial autocorrelations appear to cut off after lag 3, indicating AR(3) behaviour. Tentatively, the fitting models were ARIMA (3,1,0) and ARIMA(0,1,1) for undamaged case to forecast any movement of the bridge. These models also included with the constant term to allow for the fact that the series differences appears to vary about a level greater than zero. If Y_t denotes the GPS data, then the differenced series is $\Delta Y_t = Y_t - Y_{t-1}$.

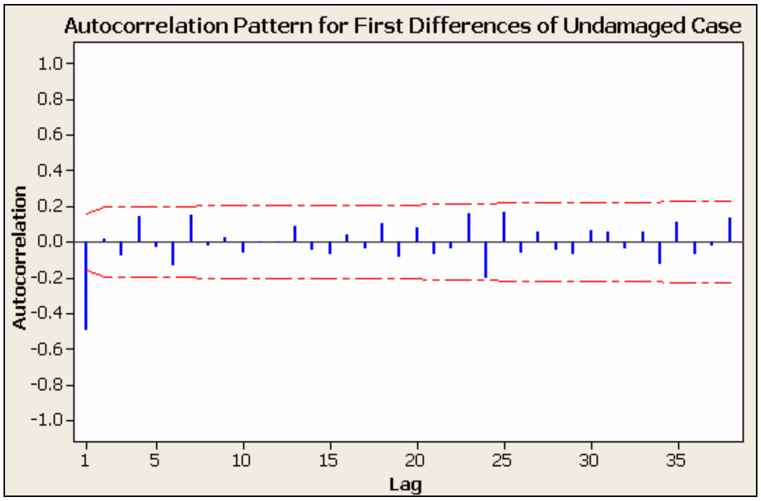

Figure 4: Autocorrelation Function for Differences

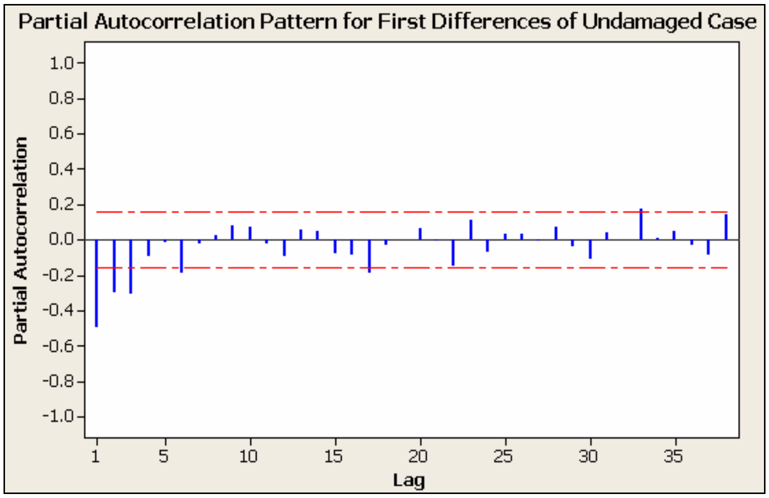

Figure 5: Partial Autocorrelation Function for Differences

3.2 Results and Discussion

The Minitab outputs for ARIMA(3,1,0) models are shown in Table 3 and the Minitab outputs for ARIMA(0,1,1) are shown in Table 4. Result for the ARIMA(3,1,0) model can referred to Figure 6 and Figure 7 while result for ARIMA(0,1,1) model referred to Figure 8 and Figure 9. Both of the models are shown with residual plot which are contains a histogram and normal probability to check for normality and observation order to check for outliers.

Table 3: Minitab output for ARIMA(3,1,0) model

```
Final Estimates of Parameters

Type            Coef    SE Coef       T      P
AR   1       -0.7842     0.0765  -10.25  0.000
AR   2       -0.5463     0.0889   -6.14  0.000
AR   3       -0.3305     0.0766   -4.32  0.000
Constant  -0.0000047  0.0001414   -0.03  0.973


Differencing: 1 regular difference
Number of observations:  Original series 154, after differencing 153
Residuals:    SS =  0.000455558 (backforecasts excluded)
              MS =  0.000003057  DF = 149


Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag            12     24     36     48
Chi-Square   15.0   28.6   50.0   68.2
DF              8     20     32     44
P-Value     0.060  0.097  0.022  0.011


Forecasts from period 154

                      95% Limits
Period  Forecast     Lower     Upper  Actual
   155  0.326665  0.323237  0.330093
   156  0.326830  0.323323  0.330337
   157  0.327095  0.323455  0.330735
   158  0.327234  0.323424  0.331044
```

Table 4: Minitab output for ARIMA(0,1,1) model

```
Final Estimates of Parameters

Type            Coef     SE Coef      T      P
MA   1        0.8360      0.0455  18.36  0.000
Constant  0.00001186  0.00002313   0.51  0.609


Differencing: 1 regular difference
Number of observations:  Original series 154, after differencing 153
Residuals:    SS =  0.000438751 (backforecasts excluded)
              MS =  0.000002906  DF = 151


Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag            12     24     36     48
Chi-Square   11.9   27.2   42.5   59.9
DF             10     22     34     46
P-Value     0.290  0.205  0.150  0.082


Forecasts from period 154

                      95% Limits
Period  Forecast     Lower     Upper  Actual
   155  0.328029  0.324688  0.331371
   156  0.328041  0.324655  0.331427
   157  0.328053  0.324623  0.331483
   158  0.328065  0.324591  0.331539
```

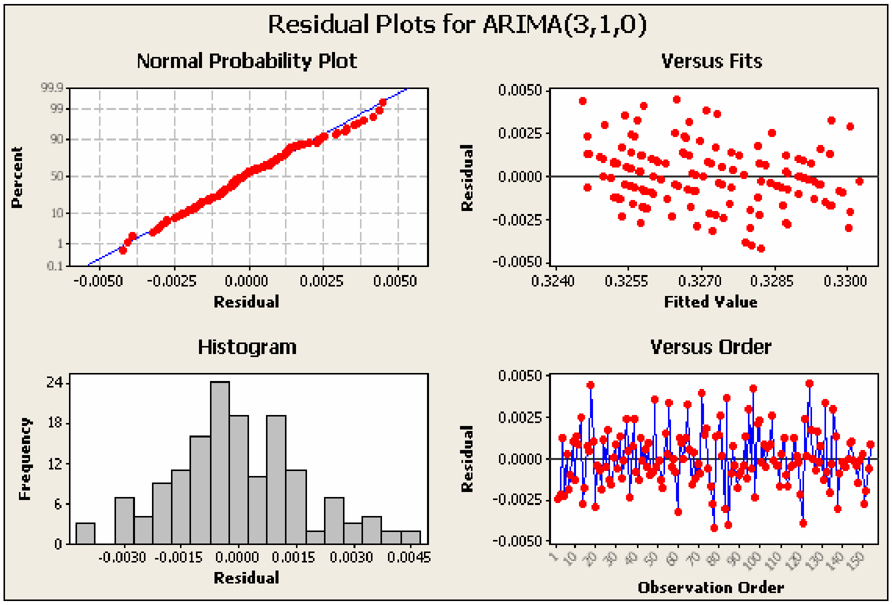

Figure 6: Residuals Plot for ARIMA(3,1,0) model

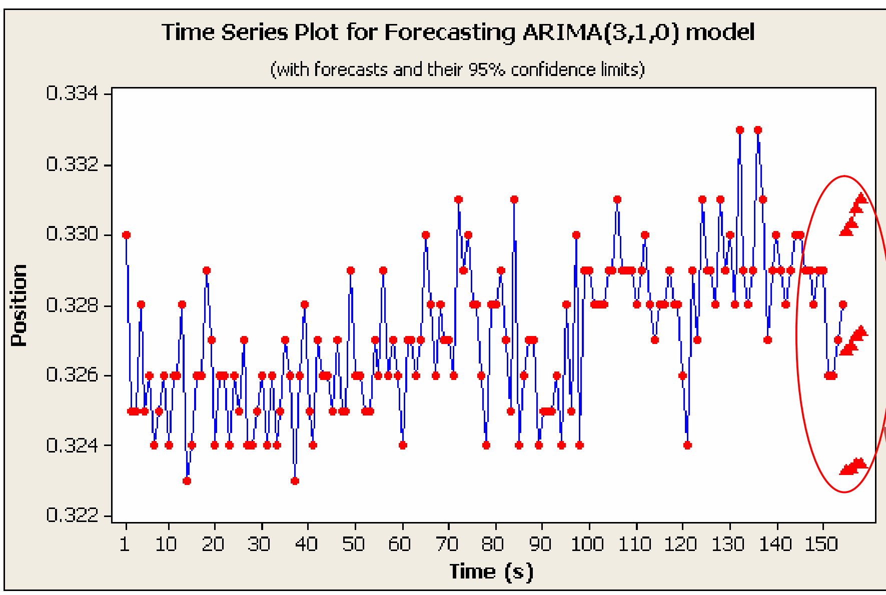

Figure 7: Forecasting Plot for ARIMA(3,1,0) model

ARIMA(3,1,0) show the four points of future values for forecasting

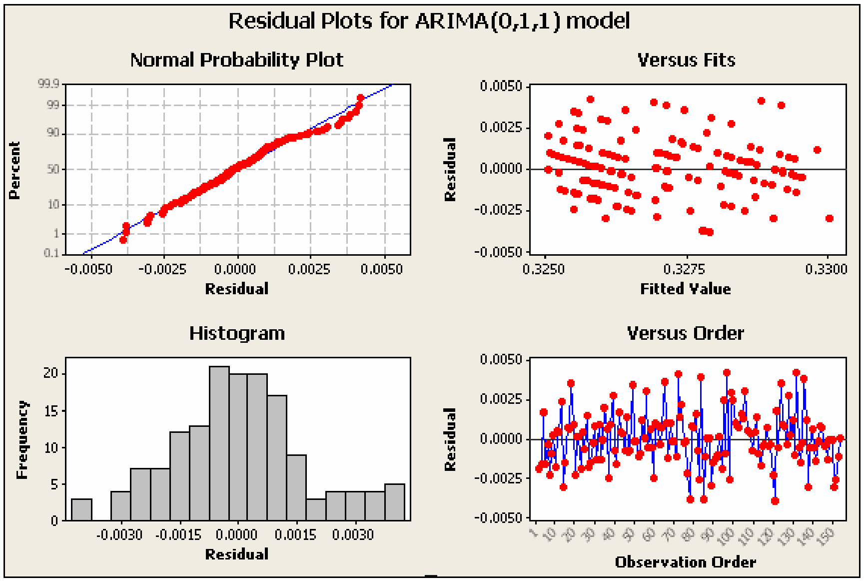

Figure 8: Residual Plot for ARIMA(0,1,1) model

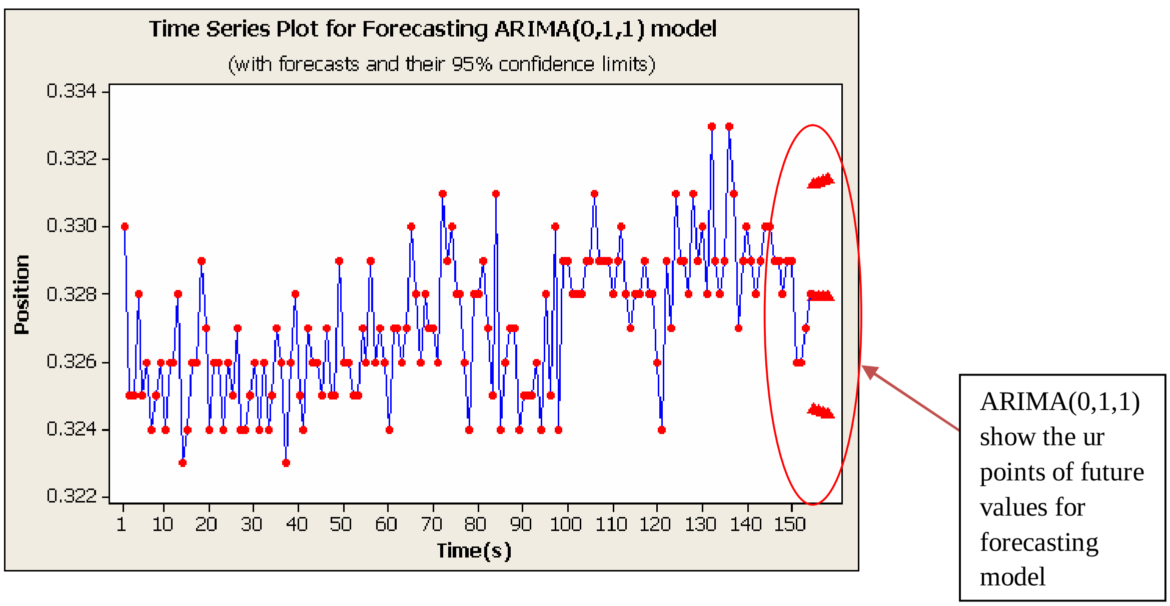

Figure 9: Forecasting plot for ARIMA(0,1,1) model

Also observation plot vs. residuals to check for outliers. Both models appear to fit the data well. The mean square errors are similar:

ARIMA(3,1,0) : S2 = 0.000003057

ARIMA(0,1,1) : S2 = 0.000002906

ARIMA models are identified (selected) by looking at a plot of the series and by matching sample autocorrelation and sample partial autocorrelation patterns with the known theoretical patterns ARIMA processes. However, there is some subjectivity involved in this procedure, and it is possible that two (or more) initials models may be consistent with the patterns of the sample autocorrelation and partial autocorrelations. Moreover, after estimation and checking, both models may adequately represent the data but by looking to the number of parameters, the principle Parcimony was used to select the simpler model or select the smallest mean square error is ordinarily preferred.

As the result, models ARIMA(0,1,1) were used as a forecasting model to detect any movement of the structure. According to Figure 9, look at end of the model, there have 4 points of forecasting data develop resulted from the Minitab software. For every point of forecasting data, it contains forecast, lower and upper limits. In the real application, any GPS data for structural monitoring exceeded to the limits of the range can considered as an unusual event happened to the structural. For safety, quick inspection should be done.

4. Conclusion

This paper has successfully examined and demonstrated the application of concept of time series analysis to the processing of data from a continuously operating Structural Health Monitoring system installed to the simulation Vibration Bridge. The ARIMA model can be used as model forecasting to detect the movement of the bridge using the GPS data. For future work, more tests will be carried out by employing the developed ARIMA model on real structures and it is anticipated that the current trend of prediction could be achieved and better results could be obtained too.

References

Chang F. K., (2001). *Proceedings of the First, Second and Third International Workshops on Structural Health Monitoring,* Stanford University, CA, CRC Press, New York.

Straser E. G. And A.S. Kiremidjian, A.S., (1998). Modular wireless damage monitoring system for structures, *Report No. 128, John A. Blume Earthquake Engineering Center,* Department of Civil and Environment Engineering , Stanford University, Stanford, CA.

Lynch J.P., Sundarajan A., Law K.H., Kiremidjian A.S. and Carryer E. (2004). *Embedding damage detection algorithms in wireless sensing unit for attainment of operational power efficiency,* Smart Materials and Structures 13 (4) 800-810.

Hanke J. E., and. Wichern D. W, (2005). *Business Forecasting,* eight ed., Prentice Hall, New Jersey,

Box, G.E. P., Jenkins, G.M. and Reinsel, G.C. (1994). *Time Series Analysis: Forecasting and Control,* Third Edition. Upper Saddle River, NJ: Prentice Hall.

Doebling *et al*. (1996). *Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics: a literature review*, Los Alamos National Laboratory Report LA-13070-MS, Los Alamos National Laboratory, Los Alamos, NM 87545, 1996.

Sohn, H., Farrar C.R., Hunter H.F. and Worden, K. (2001). Applying the LANL statistical pattern recognition paradigm for structural health monitoring to data from a surface-effect fast patrol boat, Los Alamos National Laboratory Report LA-13761-MS, Los Alamos National Laboratory, Los Alamos, NM 87545,.

Lei Y., Kiremidjian A.S., Nair, K.K., Lynch J.P., Law K.H., Kenny T.W., Carryer E. And Kottapalli A. (2003). *Statistical damage detection using time series analysis on a structural health monitoring benchmark problem*, Proceedings of Ninth International Conference on Application of Statistics and Probability in Civil Engineering, ICASP, San Francisco, CA.