

## ADJUSTMENT OF CADASTRAL NETWORK USING LEAST-SQUARES VARIANCE COMPONENT ESTIMATION

N. K. Bidi<sup>1\*</sup>, A. H. M. Din<sup>1,2\*</sup>, Z. A. M. Som<sup>1</sup> and A. H. Omar<sup>1</sup>

<sup>1</sup>Geomatic Innovation Research Group (GIG), Faculty of Built Environment and Surveying,  
Universiti Teknologi Malaysia, Johor, Malaysia.

<sup>2</sup>Geoscience and Digital Earth Centre (INSTEG), Faculty of Built Environment and Surveying,  
Universiti Teknologi Malaysia, Johor, Malaysia.  
nkhalilah3@yahoo.com; amihassan@utm.my

**KEY WORDS:** Least Squares Adjustment; Variance Component Estimation; LS-VCE; Cadastral Network; Chain; EDM; Total Station

### ABSTRACT:

The role of the stochastic model very important in data processing of geodetic network since it describes the accuracy of the measurements and their correlation with each other. Knowledge of weights of the observables is required to provide a better understanding of the sources of errors and to model the error, hence the weights need to be determined correctly. This study concentrates on the estimation of variance components from different types of instruments used in the cadastral survey. The ideas are to combine the conventional and advanced instruments in a traverse network to enhance the estimated variance component in the stochastic model. Thus, Least Squares Variance Component Estimation (LS-VCE) method was used in this study because the method is simple, flexible and attractive due to the precision of variance estimators that can be directly obtained. Observation data come with several types of instruments such as chain measurement, Electronic Distance Measurement and total station were utilized. The findings showed that LS-VCE method was very reliable in cadastral network application. Besides, the results revealed that the estimated variance components for distance scale error,  $\sigma_p$ , seem to become unrealistic for each data tested. It was found that the traverse network which included chain survey, showed the insignificant result to the precision of station coordinates when the measurements were combined.

### 1. INTRODUCTION

Data processing in geodetic network often count on the least squares adjustment (LSA) method, thus an appropriate stochastic model of the observables is constantly required (Amiri-Simkooei, 2007). In addition, information of weight of the observables is important to estimate the correct least squares solution (Wentworth, 1965). Inappropriate weight (variance component) can affect the LSA results, which is significant in geodetic network. According to Zhou (2017), different choices of weights denote different approximation methods which can affect to unbiased estimates with different statistical level efficiency. The input data of the geodetic network representing the measurement accuracy is also referred to as the variance components of the measurement data. Therefore, the estimation of the unknown variance components is called variance component estimation (VCE).

In recent years, conventional approaches have been used to estimate the variance component. Such methods that used include employing a rigorous computation, adopting the accuracy of specification given by the manufacturer, based on analysis of previous works, using individual groups of the observations for separating adjustments of the network, and last but not least, trial and error method, in which different combinations of the suspected variances of the observations are entered into the adjustment (Dennler, 1980; Chen, 1983). Obviously, the conventional methods are too rigid and need rigorous computation. Also, they are not suitable for current practice and with current needs such as new technologies like GPS, UAV, datum unification, photogrammetry, etc.

Since the conventional methods have their own weakness, the geodetic community has been trying to explore a better alternative. As a result, several approaches based on statistical methods have shown the potential to be more convincing. At present, many different methods exist for estimating the variance component (VCE). One of the well-known methods is the Minimum Norm Quadratic Unbiased Estimator (MINQUE). The other methods are the Best Invariant Quadratic Unbiased Estimator (BIQUE), the Restricted Maximum Likelihood Estimator (REML), the Bayesian method, and Least Squares Variance Component Estimator (LS-VCE) to the VCE. The VCE approach is being proposed to be implemented in the attempt to develop more realistic weighting factors for geodetic network adjustment. The principle of VCE method generally is based on the unconstrained estimation (Amiri-Simkooei, 2016).

Regarding the previously stated scenario, the conventional methods are not challenged enough and cannot be achieved by the following problem statement. The determination of standard deviation (or the estimation of variance component) is required for geodetic networks. The present practice of determining the variance component through conventional means is faced with difficulties. Thus, the geodetic community is actively seeking an alternative method for the determination of VCE.

To address this challenge, Amiri-Simkooei (2007) come out with a new method called LS-VCE, where using the principle of least square and weight matrix is defined by the user. The main idea of this method is to identify whether it is possible to apply the least squares estimation method and solve VCE

problems. LS-VCE method also sustains all the basic properties of a least squares estimator. The advantages of these methods are that the existing knowledge of least squares theory can be applied for various aspects of VCE and the precision of variance components can directly be achieved (Teunissen and Amiri-Simkooei, 2008). Therefore, this paper will highlight and focus on the estimation of variance components from different types of data by using LS-VCE method for the cadastral network survey.

For many applications, information available on the covariance matrix of the observation is crucial. This is because every measurement data or observation data of geodetic network are necessary to know their level of accuracy in the form of a standard deviation or variance. Thus, the standard deviation of the measurement has supposedly become one of the important elements in LSA. A realistic covariance matrix allows us to determine a correct precision description of the unknown parameters. For geodetic applications, it is crucial to have knowledge about the covariance matrix of the observables since that variance component is the most commonly used to determine a realistic precision. In addition, it allows us to investigate the various contributing error factors in the observations (Amiri-Simkooei, 2007). Such an example in Figure 1, what if a set of measurement data is computed by using StarNet software have failed the Global test (Chi-Square test). However, by removing blunders in observations or use trial and error method by changing the standard deviation values seem not practical to be done in bundles of data. Thus, determination of variance component in more systematic ways are crucially needed. This shows the importance of variance component estimation methods.

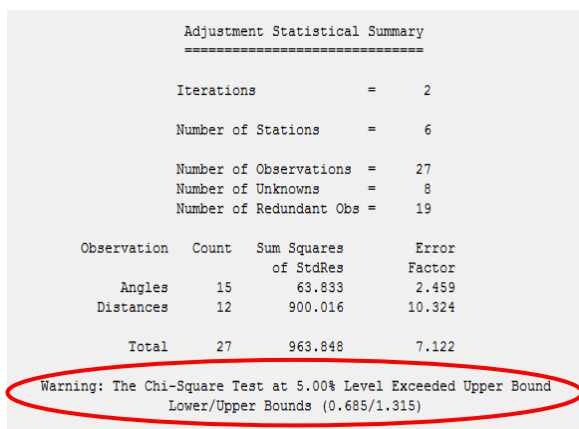


Figure 1. Example of a set of measurement data have failed Chi-Square test in StarNet software.

### 1.1 VCE to Cadastral Network Adjustment

Over the past decades, cadastral survey has undergone many technical, operational, structural and institutional changes (Yusoff *et al.*, 2013). Advancement of technologies and the enhancements in the measurement techniques have made the cadastral survey a primary task in geodetic applications. Moreover, the precision and accuracy of the measurements have become much better. Currently, there are many modern survey equipment available such as theodolite, Electronic Distance Measurement (EDM) and total station to surveyors to replace the outdated land survey equipment. Nowadays, the total station is one of the instruments that is mainly used for determining the land area information (Kavitha *et al.*, 2018). It is well known that Bowditch method is conventional method

used for cadastral data processing where the correction of the observations is applied irrespective of their uncertainties. Nevertheless, at the present time, the least squares adjustment technique has been introduced to substitute the old Bowditch method in the distribution of survey errors as the ultimate proof of boundary mark position (Yusoff *et al.*, 2013).

In addition, the least squares adjustment method has gained popularity as the availability of computers has made it possible to develop the calculations routine. More importantly, the least square adjustment is valid for any type of traverse and has the benefit that observations of varying precisions can be weighted properly in the computations (Yusoff *et al.*, 2013). In Malaysia, MicroSurvey StarNet software is commonly used among surveyors in order to process data related to cadastral survey. StarNet software able to handles general purpose like analysis of rigorous least squares and adjustments of 1D, 2D and 3D land survey networks (Starplus, 2000; Olyazadeh *et al.*, 2011).

According to Bozic *et al.* (2011), selection of instruments and measurement methods are usually related to the stochastic model of a geodetic survey. The instrument's technical specifications give the data on the instruments, and they provide the characteristic of precision for the measurement. The need for knowledge of variance component estimation has increased tremendously over the years but the measurement accuracy tends to degrade due to many error sources. Therefore, this paper will highlight and emphasis on the estimation of variance components from different types of data by using LS-VCE method particularly on the cadastral network.

## 2. DATA AND METHODS

The study utilizes cadastral data by using different type of instruments obtained from the Department of Survey and Mapping Johor Bahru (DSMM). These instruments include observation by chain measurement, electronic distance measurement (EDM), theodolite and total station (TS). The study area is in Johor state at Mukim Tebrau region. The data were picked randomly from the year 1966 to 2014. This is because, with the advancement of technology, various modern instruments were developed.

Two cadastral networks were tested in this study. The first network is an instrument combination of EDM, theodolite and total station. For the second network, the instruments used are chain measurement, EDM, theodolite and total station. All datasets of the cadastral network are summarised in Table 1. Figure 2 and 3 represents the following cadastral network surveying.

Dataset	Instruments	Stations
1	EDM Theodolite Total station	11
2	Chain EDM Theodolite Total station	18

Table 1. Combination of the instruments used for datasets in cadastral network

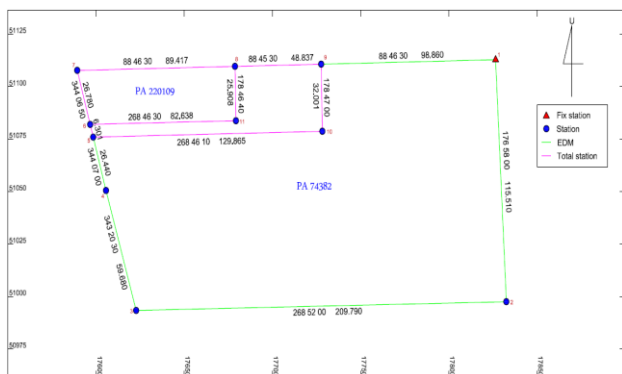


Figure 2. Data 2: Lot combination from electronic distance measurement (EDM) and total station (TS)

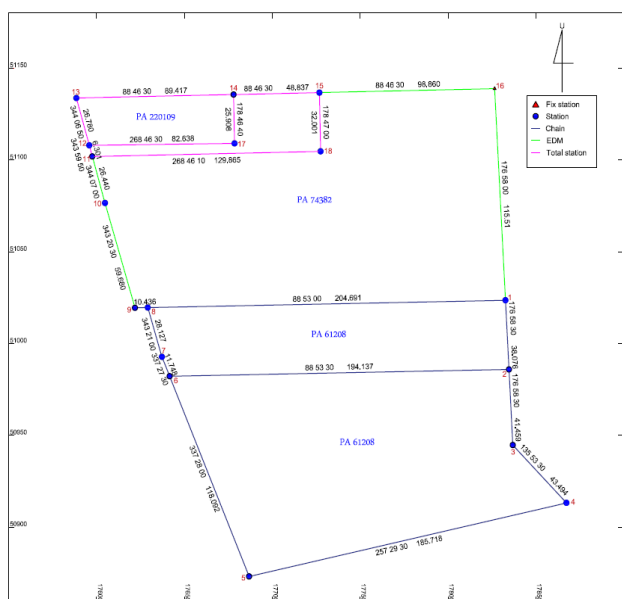


Figure 3. Dataset 3: Lot combination from chain measurement, EDM and total station

### 2.1 Procedure of LSA Computation

Data processing and computation of this study is started with the implementation of the least squares adjustment (LSA) to the cadastral networks. LSA is a systematic and simple method to compute estimated values of parameters (i.e. unknown parameter) from redundant measurements. LSA is not required if there is no redundant observation. Figure 4 explains the basic procedures to perform the least squares adjustment.

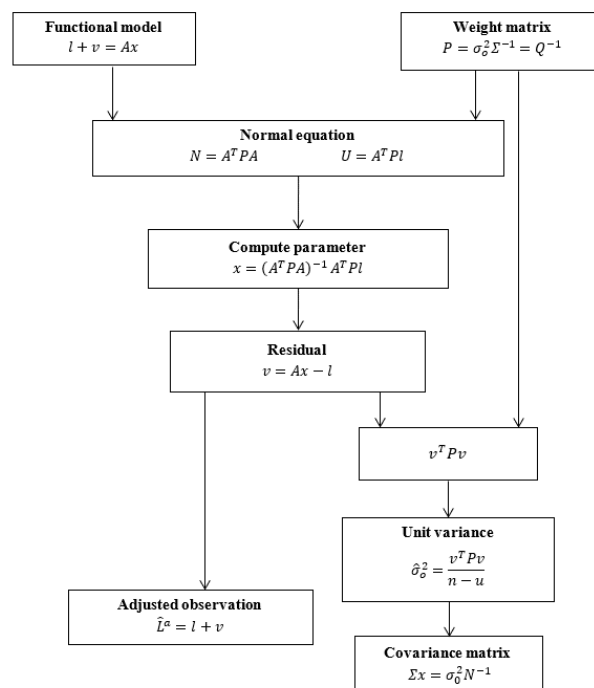


Figure 4. The procedure of least squares adjustment computation

### 2.2 Procedure of VCE Computation

After LSA computation is done, VCE computation is performed using LS-VCE method in order to compute the estimators  $\hat{\sigma}_k$ . However, as the variance components  $\sigma_k$  are unknown a priori, iterations could be performed. It could start with an initial guess for the  $\sigma_k$ . From these values, the computation could be performed by using the equation suggested by Amiri-Simkooei (2007) to estimate for the  $\sigma_k$ , which in the next iteration are considered the improved values for  $\sigma_k$ . The procedure is repeated until the estimated components do not change by further iteration. If the guess for  $\sigma_k$  equals the computed estimate  $\hat{\sigma}_k$ , the convergence of the variance components is reached. As a simple guide, for the case of the parametric adjustment model, the procedures for the computations of the variance covariance components are summarised as shown in Figure 5.

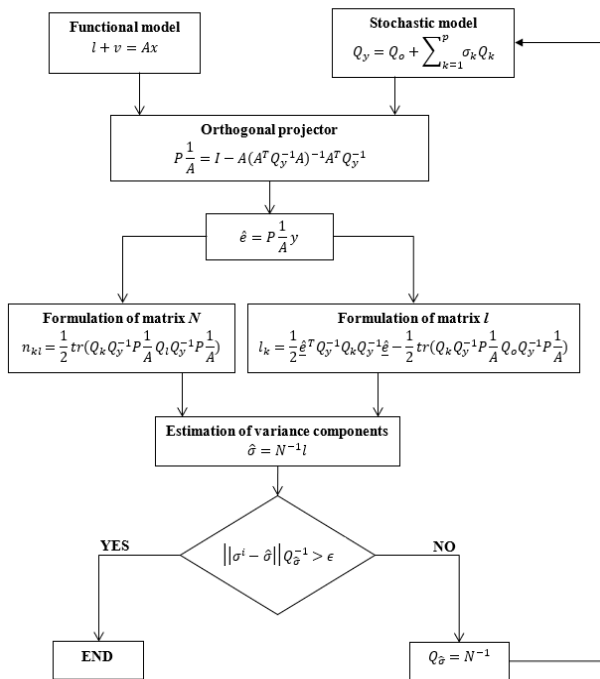


Figure 5. The computation procedure of variance covariance components using LS-VCE method.

Next, the estimated components are used as a weight matrix,  $P$  for the second adjustment of the network to see the significance of VCE approach towards the result of LSA. Statistical testing of variance components is performed using W-test in this study to identify the significance of estimated components and to select the most appropriate model for the geodetic networks. Amiri-Simkooei *et al.* (2007) stated that LS-VCE can use statistical hypothesis testing in the stochastic model same as work with the functional model over other method. The hypotheses testing can be considered as:

$$H_0: Q_y = \sum_{k=1}^p \sigma_k Q_k; \quad (1)$$

$$H_a: Q_y = \sum_{k=1}^p \sigma_k Q_k + C_y \nabla \quad (2)$$

where  $C_y$  = known cofactor matrix  
 $\nabla$  = unknown variance parameter

To test  $H_0$  against  $H_a$ , the generalized likelihood ratio test can be used. Then, W-test statistic can be achieved as follows (Amiri-Simkooei *et al.*, 2007):

$$W = \frac{b \hat{e}^T C_y \hat{e} - \text{tr}(C_y P_A^1) \hat{e}^T \hat{e}}{\sigma_w^2 [2b^2 \text{tr}(C_y P_A^1 C_y P_A^1) - 2b \text{tr}(C_y P_A^1)]^{1/2}} \quad (3)$$

where  $b$  = redundancy of functional model  
 $\hat{e}$  = least squares residual  
 $P_A^1$  = orthogonal projector

Based on the theory, the standard normal distribution of large number of observations, W-test statistic can be concluded as:

$$W \sim \sum_{i=1}^r \lambda_i \chi_i^2(1,0) \quad \text{under } H_0 \quad (4)$$

where the expectation and the variance of the W-test statistic are 0 and 1, respectively. However, the goal of W-test is to compute for different alternative hypothesis, such as difference in cofactor matrix,  $C_y$  in the preceding equation and to select the lowest value for the W-test.

### 3. RESULT AND DISCUSSIONS

#### 3.1 Dataset 1 – EDM, Theodolite and Total Station

Dataset 1 involved the traverse network that combined the measurement from EDM, theodolite and total station instruments. There were 11 stations in the traverse network with 13 distances and 13 bearing azimuths of the observation data. A total of three stochastic models were suggested for Dataset 1, where the model included *distance-dependent model* (Model 1), *distance-independent model* (Model 2) and *combined model* (Model 3). Model 1 comprised the distance-dependent component,  $\sigma_d$  and bearing component,  $\sigma_b$ . Model 2 included the distance scale component,  $\sigma_p$  and bearing component,  $\sigma_b$ , while Model 3 consists all the components which distance-dependent component,  $\sigma_d$ , distance scale component,  $\sigma_p$  and bearing component,  $\sigma_b$ . Table 2 denotes the summary of the estimated variance components for all three stochastic models as suggested in Dataset 1. The lowest variance components could be seen in Model 1, which involved  $\sigma_d$  and  $\sigma_b$  component only.

Model	Instruments	$\sigma_d$ (m)	$\sigma_p$ (ppm)	$\sigma_b$ (sec)
1	EDM/Theodolite	0.007	-	8
	TS	0.001	-	2
2	EDM/Theodolite	-	688	11
	TS	-	30	3
3	EDM/Theodolite	0.013	1146	13
	TS	0.001	144	3

Table 2. Results of VCE for traverse network Dataset 1 using LS-VCE

Figure 6 visualizes the standard deviation results of station coordinate for traverse network Dataset 1 which combine the measurement from EDM, theodolite and total station. As seen in Figure 6, Model 1 indicates the lower standard deviation of station coordinates after readjustment was applied by using estimated variance components. Whereas, Model 2 and 3 show the precision of station coordinates become worse in the second adjustment compared to the first adjustment. However, in order to select the most appropriate model on this traverse network, W-test statistic must be performed on the variance components. The results of W-test statistic are given in Table 3.

Model	Inst.	Initial W-test			Final W-test		
		$w_1$	$w_2$	$w_3$	$w_1$	$w_2$	$w_3$
1	EDM/ Theodolite	-0.01	-	43.4 9	-0.01	-	56.96
	TS	-0.05	-	0.42	-0.07	-	1.06
2	EDM/ Theodolite	-	0.01	1.63	-	0.03	1.44
	TS	-	-0.01	0.15	-	-0.02	0.28
3	EDM/ Theodolite	0.01	0.01	0.75	<b>0.01</b>	<b>0.01</b>	<b>0.48</b>
	TS	-0.01	-0.004	0.08	<b>-0.01</b>	<b>-0.004</b>	<b>0.02</b>

$w_1$  = distance-dependent;  $w_2$  = distance-scale;  $w_3$  = bearing azimuth

Table 3. Results of W-test for Dataset 1 using LS-VCE

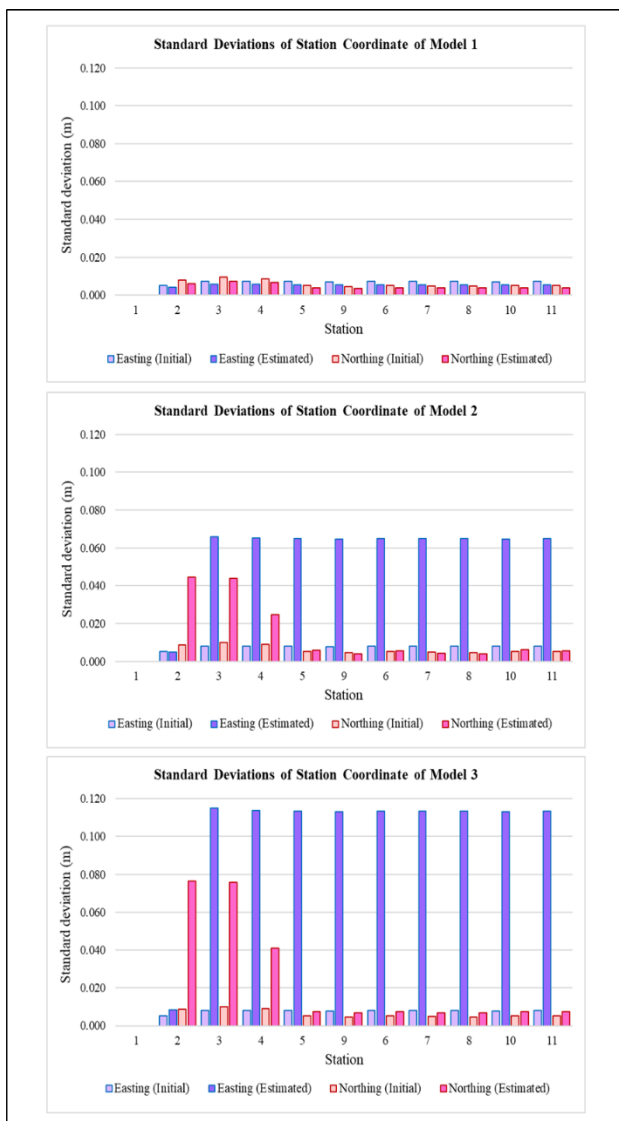


Figure 6. Estimated standard deviation of each station for Dataset 1 (EDM, theodolite and total station)

Based on the results showed, Model 3 which was the combined model indicated the smallest absolute values of W-test statistic. This means that Model 3 has less error contribution in its adjustment computation. Even though the

model showed the lowest W-test values, but the distance scale component,  $\sigma_p$  was not significant to the traverse network. Thus, Model 1 is selected as the best stochastic model on the traverse network for Dataset 1. This is because the result from LSA present the better precision of station coordinates compared to other models. In short, it could be said that the combination of EDM and total station instruments in a traverse cadastral network might be successful in this study. This also can improve the weight at covariance matrix since the variance component is the most commonly used to determine a realistic precision.

### 3.2 Dataset 2 – Chain, EDM, Theodolite and Total Station

Traverse network of Dataset 2 involves the combination measurements that use chain instrument, theodolite, EDM and total station. The observations data of Dataset 2 comprises 25 distances and 25 bearing azimuths. Dataset 3 also indicates three stochastic models which are the *distance-dependent model* (Model 1), *distance-independent model* (Model 2) and *combined model* (Model 3). The components of the stochastic models are similar as mentioned in Dataset 1 above.

Table 4 indicates the results of variance component estimation for traverse network Dataset 2 which combine the measurements from chain, EDM, theodolite and total station. From Table 4, Model 1 shows the minimum standard deviation in terms of distance-dependent component compared to Model 2 and Model 3. However, when it came to additional distance scalar component, the results show the estimated standard deviation are quite big and would be considered insignificant to the geodetic network.

Model	Instruments	$\sigma_{dl}$ (m)	$\sigma_p$ (ppm)	$\sigma_b$ (sec)
1	Chain/Theodolite	1.420	-	308
	EDM/Theodolite	0.337	-	308
	TS	0.052	-	98
2	Chain/Theodolite	-	7743	306
	EDM/Theodolite	-	156249	306
	TS	-	7711	100
3	Chain/Theodolite	0.959	20149	306
	EDM/Theodolite	1.507	270165	306
	TS	0.052	974	102

Table 4. Results of VCE for traverse network Dataset 2 using LS-VCE

Figure 7 clearly visualizes the results of the standard deviation for the first and second adjustments by employing LS-VCE method. Based on the results in Figure 7, mostly the standard deviation of stations for the second adjustment is larger than the first adjustment. Besides, the results of W-test statistic values for Model 1, 2 and 3 are depicted in Table 5.

From Table 5, the finding shows clearly that Model 1 has the smallest absolute values in terms of  $w_1$  and  $w_3$  on the final result of W-test statistic (values on the right). Complementary to choose the best stochastic model in Dataset 2, the Model 1 illustrates the better standard deviation in Figure 7, where it less than 2meter accuracy compared to some of the stations in Model 2 and 3 are more than 8 meter and above. Thus, Model 1 is the most suitable stochastic model for Dataset 2

corresponding their characteristics in precision of station coordinates and the smallest W-test values. Although Model 1 is sufficient with the characteristics, some of the stations showed the higher standard deviation after readjustment. Hence, it could be said that the VCE approach was not significant to the traverse network of Dataset 2 which involves the measurement that combined chain, EDM, theodolite and total station instruments.

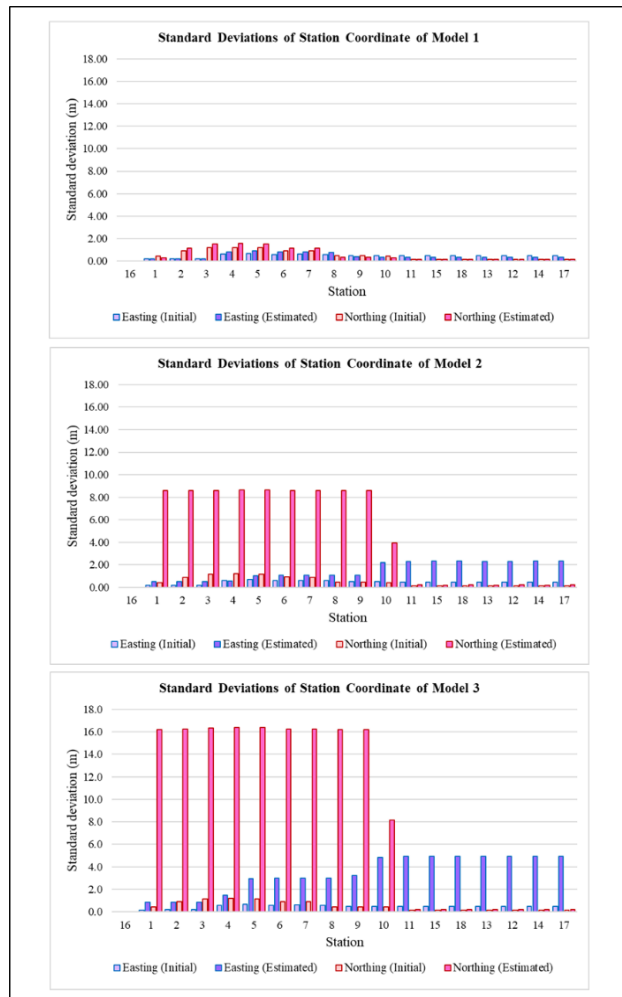


Figure 7. Estimated standard deviation of each station for Dataset 2 (chain, EDM, theodolite and total station)

Model	Instruments	Initial W-test			Final W-test		
		$w_1$	$w_2$	$w_3$	$w_1$	$w_2$	$w_3$
1	Chain/Theodolite	-174.14	-	610.22	<b>-104.89</b>	-	<b>363.11</b>
	EDM/Theodolite	-159.46	-	610.22	<b>-94.00</b>	-	<b>363.11</b>
	TS	-189.61	-	-114.75	<b>-115.54</b>	-	<b>-61.64</b>
2	Chain/Theodolite	-	-13.91	59.13	-	-380.54	1618.85
	EDM/Theodolite	-	-14.49	59.13	-	-397.12	1618.85
	TS	-	-14.77	-4.27	-	-404.52	-119.41
3	Chain/Theodolite	-16.32	-13.87	59.02	-408.63	-347.57	1490.06
	EDM/Theodolite	-16.13	-14.45	59.02	-405.33	-363.32	1490.06
	TS	-18.44	-14.74	-4.26	-463.04	-369.77	-98.17

$w_1$  = distance-dependent;  $w_2$  = distance-scale;  $w_3$  = bearing azimuth

Table 5. Results of W-test for Dataset 2 using LS-VCE

To summarise, based on the overall results and analysis of the tested data on the cadastral network of this study, it is found that the traverse network that included chain survey showed insignificant result to the standard deviation of station coordinates. However, when it came to traverse network that used more accurate instrument such as EDM and total station, the results showed the realistic standard deviation in tandem with or closer to the accuracy given by manufacturers. Therefore, it can be summarized that the combined measurements, which included chain measurement, neither least squares adjustment (LSA) or estimated variance components (VCE) can be performed. However, the LSA computation can be performed if there was only contained the chain measurements in a traverse network.

#### 4. CONCLUSIONS

The estimated variance component of the cadastral network using least squares variance component estimation (LS-VCE) method has been presented in this paper. Cadastral network utilized the observation data from several types of instrument used such as chain measurement, EDM and total station. To summarize, the network that included chain instrument gave the worst precision and affect the results particularly the adjusted coordinates. Thus, a combination of chain measurement with other instruments was unsuccessful in this study. This contrasts with a combination of EDM, theodolite and total station instrument, it showed a new finding from this combination to this study. Hence, the combination of these instruments can be applied in the cadastral survey where it can improve the accuracy of the traverse networks and can combine more cadastral lots of different years in tandem with the advancement of instruments used.

#### ACKNOWLEDGEMENTS

The author would like to thanks to the DSMM Johor Bahru for providing cadastral data. We are grateful to the Ministry of Higher Education Malaysia and Universiti Teknologi Malaysia for partly funding this research under Research University Grant (Vot number: Q.J130000.2527.19H83 and Q.J130000.2527.19H26).

#### REFERENCES

- Amiri-Simkoei, A. R., 2016. Non-negative least-squares variance component estimation with application to GPS time series. *Journal of Geodesy*. Springer Berlin Heidelberg, 90(5), pp. 451–466.
- Amiri-Simkoei, A. R., 2007. *Least-Squares Variance Component Estimation: Theory and GPS Applications*. PhD Thesis, Delft University of Technology, Delft, Netherlands.
- Amiri-Simkoei, A. R., Teunissen, P. J. G. and Tiberius, C. C. J. M., 2007. Assessment of noise in GPS coordinate time series: Methodology and results. *Journal of Geophysical Research*, 112(7), pp. 1–19. doi: 10.1029/2006JB004913.
- Bozic, B., Gospavic, Z. and Milosavljevic, Z., 2011. Estimation of the variance components in various covariance matrix structures. *Survey Review*, 43(323), pp. 653–662. doi: 10.1179/003962611X13117748892434.

- Chen, Y. Q., 1983. *Analysis of Deformation Surveys – A Generalized Method*. Technical Report No. 94. Department of Geodesy and Geomatics Engineering, University of New Brunswick, Fredericton, N.B.
- Denkler, S. M., 1980. *Evaluation of Micro-Geodetic Networks for Monitoring Tectonic Movements in Peru*. M.Eng. Report, Dept. of Surveying Engineering, University of New Brunswick, Fredericton, N. B., Canada.
- Kavitha, M., Viswanath, R., Kavibharathi, P., Aakash, K. and Balajimanikandan, M., 2018. A Comparative Study of Conventional Surveying Techniques with Total Station and GPS. *Intern. Journal of Civil Eng. and Tech.*, 9(1), pp. 440–446.
- Olyazadeh, R., Setan, H. and Fouladinejad, N., 2011. Network adjustment program using MATLAB. *Geospatial World Forum*, (January).
- Starplus Software, 2000. *STARNET V6 Least Squares Survey Adjustment Program. Reference Manual*.
- Teunissen, P. J. G. and Amiri-Simkooei, A. R., 2008. Least-squares variance component estimation. *Journal of Geodesy*, 82(2), pp. 65–82. doi: 10.1007/s00190-007-0157-x.
- Wentworth, W. E., 1965. Rigorous least squares adjustment. *Journal of Chemical Education*, 42(2), pp. 96–103.
- Yusoff, M. Y., Jamil, H., Halim, N. Z., Yusof, N. A. and Zain, M. A., 2013. Ekadaster: A Learning Experience for Malaysia. *FIG Pacific Small Island Developing States Symposium*, 18-20 September 2013, Suva, Fiji.
- Zhou, X., 2017. A unified framework for variance component estimation with summary statistics in genome-wide association studies. *Ann. Appl. Stat.*, 11(4), pp. 2027–2051. doi:10.1214/17-AOAS1052.

Revised August 2019