Comparative Analysis of Google Earth Derived Elevation with In-Situ Total Station Method for Engineering Constructions

Njike CHIGBU[1], Maduabughichi OKEZIE[1] and Ikenna Donald ARUNGWA[2]
Chima OGBA[3] Nigeria

Keywords: Google Earth (G.E.) elevation; Total Station Levelling; Engineering Construction; Comparative Analysis; Longitudinal Profile.

Summary

Fundamental to any development on the earth surface is accurate topographic information – elevation dataset in particular. Consequently, lasting, sustainable, and enduring physical development cannot be achieved in the absence of accurately determined height/elevation dataset. Hopes are raised with the advent of satellite derived Digital Elevation datasets or models of which Google Earth (G.E.) derived elevation is part of. Accuracy of such dataset is therefore of major concern for potential users. In this paper, the accuracy of Google Earth (G.E.) derived elevation is assessed, using a 10.16 km profile elevations, obtained by means of total station levelling technique, as reference or benchmark within Aba metropolis in Nigeria. Cursory accuracy statistics reports a Mean Error of 1.65m, Root-Mean-Square Error of 2.79m, Standard Deviation of 2.27m, Median Absolute Deviation of 1.72m, and Pearson's ,Spearman's &Kendall's taucorrelation values of 0.898, 0.878&0.705 respectively. Although these initial accuracy and similarity indices suggest that G.E. elevations are useful, unfortunately further incisive statistical test like the Mann-Whitney U Test of group and the t-Test suggest otherwise. The G.E. derived elevations failed to meet up with the 0.024√K and 0.1√Kfor ordinary and rough levelling (respectively)basicrequirements/standard.On the strength of the foregoing, G.E. elevations are declared not suitable for any form of levelling operation that would eventually lead to engineering construction.

1.0 INTRODUCTION

A major ingredient of physical and infrastructural development is elevation data, also known as heights of points. They are widely used in construction of roads, rails, bridges, Dams, and other scientific applications requiring height information.

The advent of alternative sources to elevation data other than the conventional means of obtaining such data is certainly a game changer. Amidst these alternative sources of elevation is Google Earth (G. E.). One remarkable feature of this elevation data source is ease of accessibility and ready availability.

Like data sets acquired through conventional method of levelling, errors are always a part of this data. The questions that remain topical, relevant and in the minds of potential users is thus;

i. What size of errors is inherent in this data in general and in certain localities?
ii. How useful and to what extent can such data be put to use given the sizes of errors inherent in it?

It is reported by (Papasaika-Hanusch, 2012;Khalid & El-Ashmawy, 2016; Richard & Ogba, 2017; Akter, 2018) that the Shuttle Radar Topographic Mission Digital Elevation Model (SRTM DEM) formed the baseline dataset of G.E.. For all practical purposes therefore, G.E. elevation can be regarded as a Digital Elevation Model, and is consequently bound to be affected by errors sources that affect DEM.

Although the vertical accuracy of elevation data from the G.E. is not in the public domain or available to researchers, literatures are consistent in their opinion of Shuttle Radar Topographic Mission data as the baseline data used for the generation of Google Earth (G.E.) elevation dataset (ibid). Also worth mention is the fact G.E. elevation database is consistently being refined as more accurate data from other sources are made available.

A relatively few number of research has been conducted, which aims to assess the accuracy of G.E. elevation data against known benchmark data (Wang et al., 2017; Khalid & El-Ashmawy, 2016;). Literatures available to the researcher only show that G.E. data has only been evaluated in one location in Nigeria (Richard & Ogba, 2017)

Hossain(2018) evaluated G.E. data with a review of ascertain if it was a viable alternative to SRTM & ASTER. It was assessed along the lines of its similarity – in describing the topography – with SRTM & ASTER. Strong Pearson's correlation values of 0.905 & 0.88 for G.E. versus SRTM 30m & SRTM 90m were respectively reported. Also reported was Pearson's correlation value of 0.469 for G.E. versus ASTER GDEM OF 30 m. It must also be mentioned that Hossain (2018)used SRTM 30m, ASTER30m and STRM 90m resolutions as reference data for assessing the accuracy of the G.E. elevation dataset. Accordingly, standard deviations of ± 0.460m, ± 0.396m and ± 0.204m were reported for the comparison of SRTM 30m, ASTER30m and STRM 90m respectively.

In the works of (El-Ashmawy, 2016)GPS ground control points were used as reference data. G.E. data was extracted using an online tool named TerrainZonum. It was such that point data was required in a grid format that covered the area of interest and subsequently interpolated to form an elevation model. The reference data used was used to extract the height values of desired points. The research reports a Root Mean Square Error (RMSE) value of 1.85m for flat areas.

It is important to note that generation of elevation data from interpolation may introduce errors, which may arise from the interpolation techniques employed, roughness of terrains, distribution and quantity of sampled points. In addition to the above mentioned fact, when the grid size of the reference DEM is not equal with that of the test DEM, results of such comparison are likely to be biased. El-Ashmawy(2016)concluded that while G.E. Data may suffice to some engineering application, it is however inadequate to meet the standard required for fine/small scale precise engineering applications.

The research of Wang et al.(2017)reports a Mean Absolute Error (MAE), RMSE & Standard Deviation S.D. of 1.32m, 2.27m & 2.7m respectively, when G.E. data was compared with roadway elevation data. It however reported RMSE, ME, MAE & S.D. of 22.31m, 0.13m, 10.72m & 22.31m respectively when G.E. data were compared against GPS benchmark in area conterminous the United State of America. The research concludes by acknowledging that accuracy of G.E. data varies in space and that its accuracy is satisfactory along roadways.

Richard & Ogba(2017) focused majorly on comparison of DEM developed from G.E. data and that developed from DGPS data. The research reports a poor performance of G.E. DEM in representing steep slopes. Although the research focused on surface characterization ability of the G.E derived DEM, it however did not deploy robust & rigorous statistical tools in the assessment of data.

2.0 MATERIALS AND METHODS

1.1. Area of Study

The study area is at Aba metropolis in Abia State Nigeria, a low-lying land south-East of Nigeria located between longitudes **7°23'41.99'' - 7°27'32.85''E** and latitudes **5°09′11.49″-5°11′34.82″N.** The Aba metropolis has stable terrain with minimal terrain undulations.

Fig 1: The red line shows the profile path in the study area (Google Earth Pro 2019)

1.2.Test data (G.E.) acquisition.

Firstly, the set reference data was imported into Arc GIS 10.3 and converted to a point layers. It was subsequently exported to a Keyhole Markup Language (KML) file format recognizable by G.E. The exported KML files were opened in G.E. Pro environment, and add path tool of G.E. software was used to draw/trace the points along the defined profile path. Subsequently, the height of traced out points were extracted using TCX converter and exported to a Comma Separated Version (C.S.V.) file format. This file format is recognizable in Microsoft Excel software.

1.3.Data processing

All statistical analysis was performed in SPSS version 23. Data analysis performed includes, correlation analysis between reference and G.E. data, t-test of means with assumption of normality, non – parametric Mann-Whitney U test of means without normality assumption, descriptive statistics analysis leading to summary statistics such as mean, mean error, Median Absolute Deviation (MAD) range. Non-parametric correlation analysis test was also performed.

Table 1.0

S/No	software	Remark
1	ESRI ArcGIS 10.5	Used for plotting and conversion of points to KML format
2	SPSS version 23	Used for statistical analysis
3	TCX	For extraction and update of height of points
4	Google Earth Pro	Platform for obtaining G.E. elevation data
5	Microsoft Excel	For data organization and profile plotting

1.4.Accuracy assessment indicators

The accuracy assessment procedure adopted for this study is similar to those in Arungwa et al. (2018)and Wessel et al. (2018)Vertical accuracies of the G. E. Elevation dataset are assessed by comparing them to ground point whose elevations were determined by the Total station equipment. In this study also, robust statistical test & accuracy indicators were deployed.

The following statistical measures were applied to assess the error in G.E. dataset (Wessel et al (2018).

$$Mean\ Error\ (ME) = \frac{1}{N}\sum_{i=1}^{N} h_i^{G.E.} - h_i^{Ref} = \frac{1}{N}\sum_i^N \Delta h_i \tag{1}$$

$$Root\ Mean\ Square\ Error\ (\text{RMSE}) = \sqrt{\frac{1}{N}\sum_i^N (\Delta h_i)^2} \tag{2}$$

$$Standard\ Deviation\ SD = \sqrt{\frac{1}{N-1}\sum_i^N (\Delta h_i - ME)^2} \tag{3}$$

$$Median\ Absolute\ Deviation\ (MAD) = Median_j(\left|\Delta h_j - m_{\Delta h}\right|) \tag{4}$$

Where $m_{\Delta h}$ is the median (50percentile/middle quantile)

$\Delta h$is the difference calculated from the difference between corresponding height values from G.E. and total station height. Positive differences represent locations where the G.E. elevation exceeded the total station elevation; and, conversely, negative errors occur at locations where the G.E. elevation was below the total station elevation(Santillan & Makinano-Santillan, 2016; Athmania & Achour, 2010)

While the ME gives an overall idea in of the average bias in in G.E. dataset, the SD and RMSE measures G.E. dataset quality and provide insight into the distribution of deviations on either side of the mean value(Athmania & Achour, 2010).

The level of agreement between G.E. and reference dataset were evaluated using parametric and nonparametric correlation analysis.

3.0 RESULTS AND DISCUSSION

Table 2.0displays the key descriptive statistics of the heights from both data sources (total station and Google Earth). The two data shows some level of similarity in that they both report slightly varying range, mean, minimum and maximum height value. At this level of analysis there seem not to be any clear distinction between the two elevation dataset.

Table 2.0: Descriptive Statistics of heights

	N	Range	Minimum	Maximum	Mean	Std. Deviation

Total Station Height (m)	412	15.295	70.687	85.982	78.67075	5.152215
Google Earth Height	412	17	72	89	80.318	4.5646
Valid N (listwise)	412					

Table 3.0 shows the descriptive statistics of the results from the basic comparison between height from G.E. and Total station. It shows a maximum error of 8.88mand minimum of –5.29m and a mean error of 1.65m, a Root Mean Square Error (RMSE) of 2.79m and standard deviation of 2.26m.

Table 3.0 Residual (Error)		
N	Valid	412
Mean		1.647206
Std. Error of Mean		0.1116437
Median		1.592500
Std. Deviation		2.2661190
Minimum		-5.2950
Maximum		8.8880
RMSE		2.79930701

Table 4.0 Descriptive Statistics of Residual (Error)		
	N	Mean
Absolute Median Deviation	412	1.7155
Valid N (listwise)	412	

In general it can be said that G.E. data overestimates the topography of the profile by an average and maximum value of 1.65m and 8.89m respectively. The positive mean error value indicates that majority of the errors are greater than zero. Therefore G.E. height values may be said to be positively biased along the profile path.

The statistics result of a more robust statistical descriptive of the Median Absolute Deviation MAD is displayed in Table 4.0. It reports a value of 1.72m. One major advantage of this statistic is that it is “immune” to the influence of extreme error values on the central error descriptive.

Figure 2, is a display of the histogram of residual distribution. Cursory examinations of the figure suggest that the errors follow a normal distribution. However a closer examination reveals an offset-positive skew- of about 1.65m (mean) error from the reference line or zero point. This pattern simply suggests a likelihood of a systematic mean error of aforementioned size in the data.

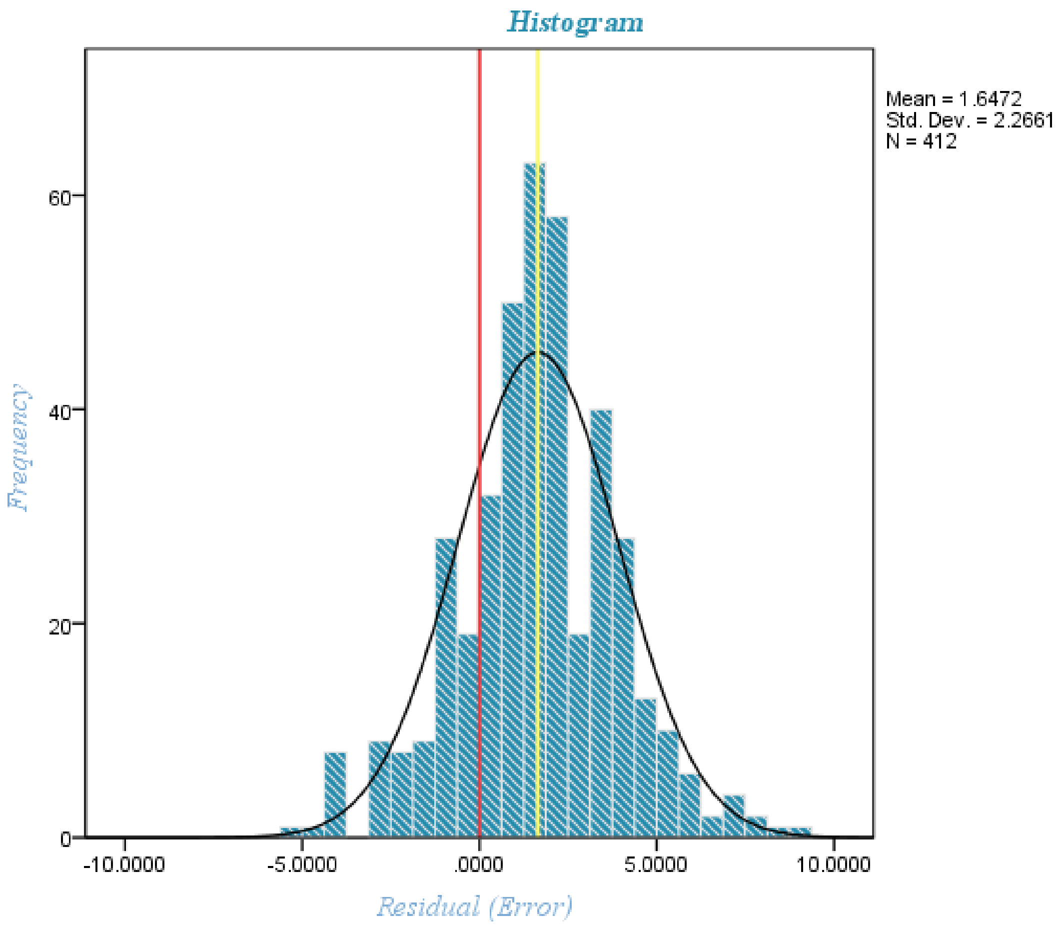


Figure 2: Histogram of residual/error distribution

One major question in the mind of prospective users of this data (Google Earth data) is thus how much similarity exists between the two datasets and how significant is this similarity. A good way to attempt such question would be to perform a correlation analysis on both datasets.

Tables 5.0, 6.0, 7.0, 8.0 and figure 3.0 were results of efforts in attempting to answer this question.The parametric correlation analysis (Tables 5.0) performed on the datasets (under the assumption of normality of data distribution), with a Pearson's value of 0.899 at 0.01 level of significance, reports the existence of a significant positive relationship between the two datasets (Total station and Google earth). This fact is further corroborated by Figure 2.

Table 5.0: Parametric Correlations Analysis from

		Total Station Height (m)	Google Earth Height	Google Minus Total Station (Residual)
Total Station Height (m)	Pearson Correlation	1	0.898**	-0.464**
	Sig. (2-tailed)		0.000	0.000

	N	412	412	412
Google Earth Height	Pearson Correlation	0.898**	1	-0.028
	Sig. (2-tailed)	0.000		0.574
	N	412	412	412
Google Minus Total Station (Residual Error)	Pearson Correlation	-0.464**	-0.028	1
	Sig. (2-tailed)	0.000	0.574	
	N	412	412	412
****. Correlation is significant at the 0.01 level (2-tailed).**				

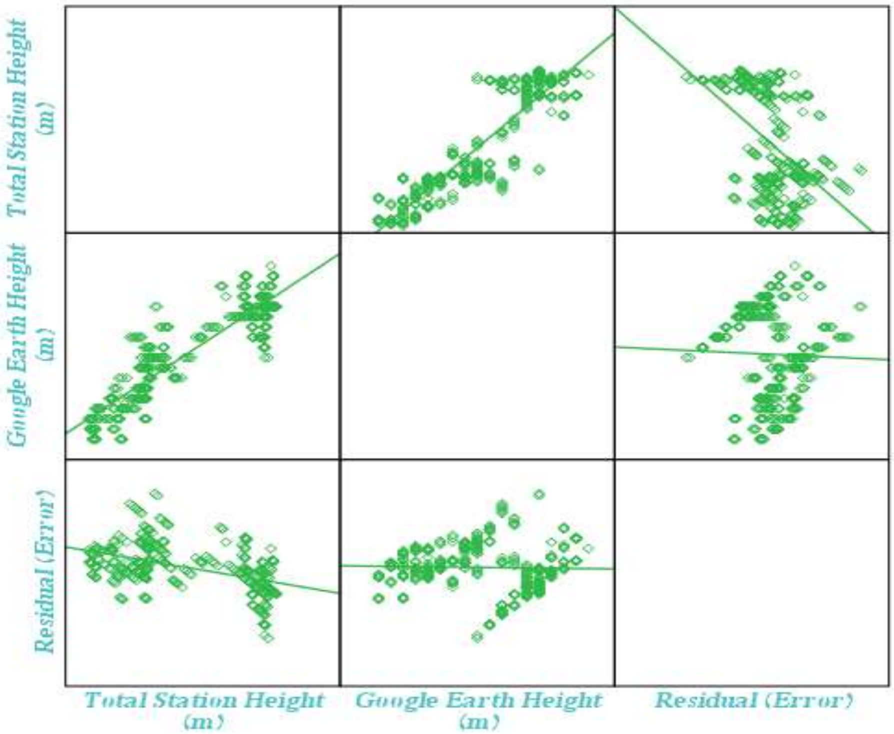


Figure 3.0: Scatter plot of G.E. vs Total Station vs Residual

With an assumption that data was not normally distributed, the datasets were subjected to a two correlation analysis-Kendall's tau and Spearman's rho; the tests, which reported respective values of 0.705 and 0.878(at 0.01 level of significance), indicate a significant relationship between datasets.

Table 6.0: Nonparametric Correlations					
			Total Station Height (m)	Google Earth Height (m)	Residual (Error)
Kendall's tau_b	Total Station Height (m)	Correlation Coefficient	1.000	0.705**	-0.301**
		Sig. (2-tailed)		0.000	0.000
		N	412	412	412
	Google Earth Height (m)	Correlation Coefficient	0.705**	1.000	0.018
		Sig. (2-tailed)	0.000		0.606
		N	412	412	412
	Residual (Error)	Correlation Coefficient	-0.301**	0.018	1.000
		Sig. (2-tailed)	0.000	0.606	
		N	412	412	412
Spearman's rho	Total Station Height (m)	Correlation Coefficient	1.000	0.878**	-0.437**
		Sig. (2-tailed)		0.000	0.000
		N	412	412	412
	Google Earth Height (m)	Correlation Coefficient	0.878**	1.000	-0.054
		Sig. (2-tailed)	0.000		0.277
		N	412	412	412
	Residual (Error)	Correlation Coefficient	-0.437**	-0.054	1.000
		Sig. (2-tailed)	0.000	0.277	
		N	412	412	412
**. Correlation is significant at the 0.01 level (2-tailed).					

In addition to the above analysis to examine similarity and significance of similarity of both datasets, the dataset was subjected to a test of means under the assumptions that data was:

1. Normally distributed
2. Not normally distributed.

The student t–tests and Mann Mann-Whitney U test of means were respectively conducted to test the similarity of the Mean. The results of both tests are contained in tables 7.0 and 8.0 respectively. Judging by ρ (Sig) value of both test which reports values way below the 0.05 benchmark, the null hypothesis that assumes that both groups and means from the two datasets are the same and equal was rejected. It is therefore safe to state that a statistically significant difference exist between the two datasets.

Another key question prospective users of the dataset for engineering works should be keen to know is thus; is there any relationship between errors/uncertainty inherent in data and altitude figures 3.0, 4.0, tables 5.0 and 6.0 gives insight to this question. Giving the ρ (Sig) values reported in tables 5.0 and 6.0with respective Pearson's, Kendall's tau and Spearman's rank correlation values of -0.464, -0.301 and -0.437, it is obvious that there exist a moderate negative statistically significant relationship between error and altitude. This is clearly revealed in figure 4.0 by the orange colour line of best fit. One can therefore safely infer that error in G.E. elevations are likely to reduce with increasing altitude/elevation of the ground within the study area.

Table 7.0: Independent Samples Test (t-Test)

Table:		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Height	Equal variances assumed	33.008	0.000	-4.857	822	0.000	-1.647206	0.339119	-2.312848	-0.981565
	Equal variances not assumed			-4.857	810.233	0.000	-1.647206	0.339119	-2.312862	-0.981551

Table 8.0:Mann-Whitney U Test of group

	Height
Mann-Whitney U	67791.000
Wilcoxon W	152869.000
Z	-5.002
Asymp. Sig. (2-tailed)	0.000

a. Grouping Variable: Height Source Code

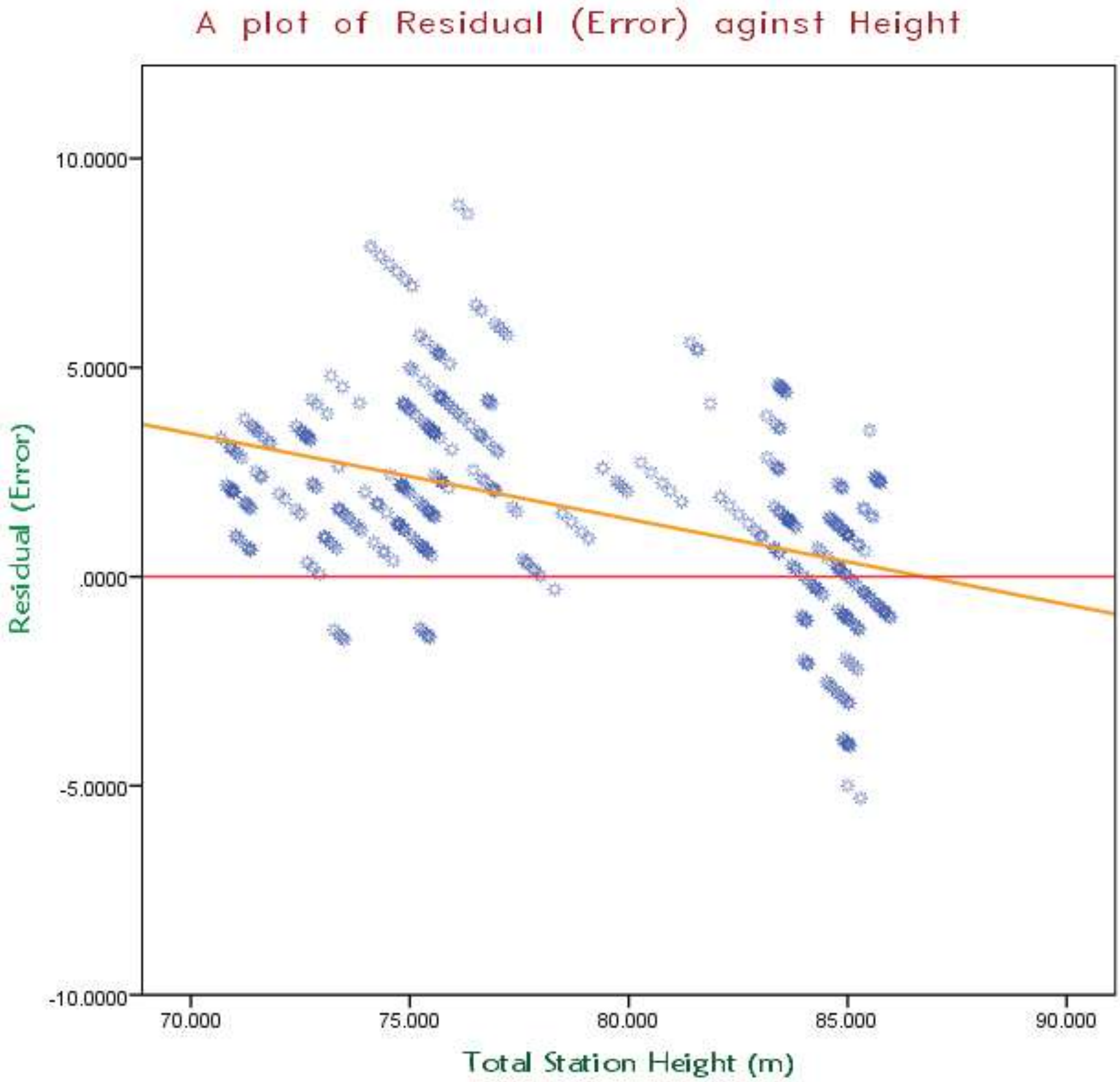


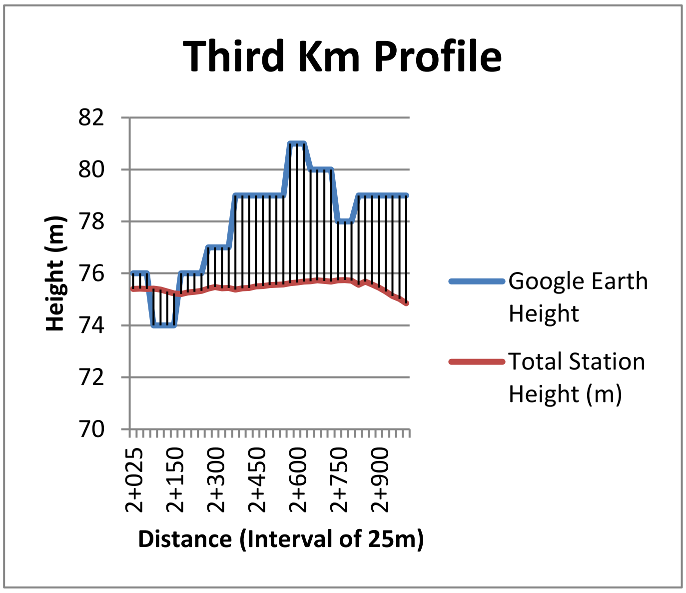
Figure 4.0: Scatter plot of residual (error) against height

Finally, describing the potential of dataset to meeting generally accepted and set standard would indeed be an apt way to either recommend or advice against itsusagefor engineering projects. The minimum (maximum closure error) error (in meters) generally accepted for levelling network is0.024√K for ordinary levelling and 0.1√Kfor rough levelling, where Kis the length of profile in Kilometres(Table9.0).

Table: 9.0 Different categories of accuracy for leveling operation

	Rough Leveling	Ordinary Leveling	Accurate Leveling	Precise Leveling	Total Length (km)
Constant value	0.1	0.024	0.01	0.005	10.125
Accuracy (m)	0.318	0.076	0.032	0.016	

Judging by the Mean Error and RMSE value of 1.65m and 2.79m**(**table 3.0) of the dataset, the G.E. height cannot be used as a sufficient replacement of heights obtained by conventional levelling method. This agrees with the works of().



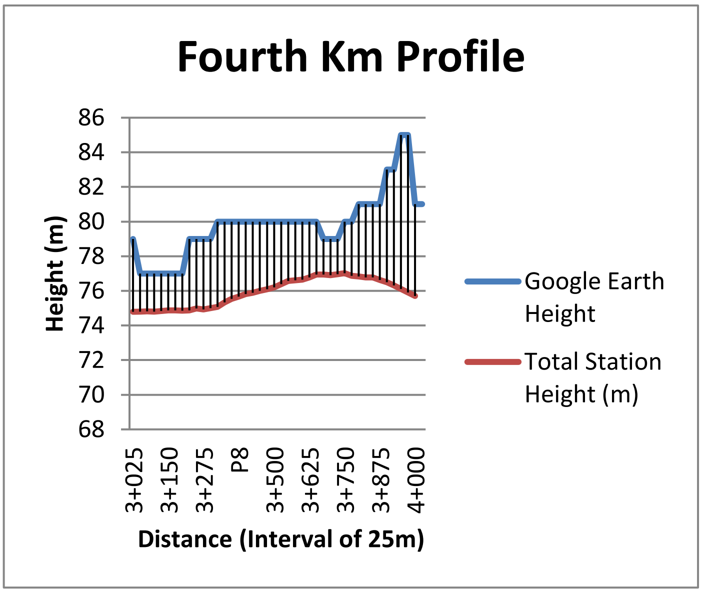


Figure 5.0: Profile of the third kilometre

Figure 6.0: Profile of the fourth kilometre

A graphic display of the performance of the dataset per kilometre is completely displayed as an appendix to this paper and an excerpt is presented in figures 5.0 and 6.0. Indeed the G.E. elevation tends to show signs of significant overestimation of the ground topography and also does not closely follow the profile of the total station data. This tendency is most undesirable by potential users.

4.0 CONCLUION

This study investigated the quality of Google Earth (G.E.) elevation data in Aba metropolis. First, the basic characteristics of G.E. elevation data were described. Then its vertical accuracy was estimated by means of comparisons against heights acquired by total station levelling along a profile path of over 10km.Finally, differences and similarity of data were assessed using statistically robust and rigorous methods. Although initial statistical test revealed significant relationship and similarity between datasets, rigorous statistical test reveals that both datasets are significantly different. In addition, the G.E. data fails to meet the minimum error standard for levelling data (table 9.0) and consequently cannot be relied or even used for serious engineering projects. It is therefore concluded that G.E. elevation dataset- within the study area- cannot be used as a comparable alternative for heights obtained via conventional levelling methods.

RFERENCES

Akter, S. (2018). Digital Elevation Modeling of Saint Martin Island, Bangladesh: A Method Based on Open Source Google Earth Data Accuracy Assessment of Google Earth Derive DEM with others Available DEM :Case Study Flat and Hilly Area of Bangladesh. View project Kinematics. *Article in International Journal of Advanced Research*, *6*(2), 379–389. https://doi.org/10.21474/IJAR01/6449

Arungwa, I. D., Obarafo, E. O., & Okolie, C. J. (2018). Validation of Global Digital Elevation Models in Lagos State, Nigeria. *Nigerian Journal of Environmental Sciences and Technology (NIJEST)*, *2*(1), 78–88. Retrieved from www.nijest.com

Athmania, D., & Achour, H. (2010). External Validation of the ASTER GDEM2, GMTED2010 and CGIAR-CSI-SRTM v4.1 Free Access Digital Elevation Models (DEMs) in Tunisia and Algeria. *Remote Sensing*, *6*, 4600–4620. https://doi.org/10.3390/rs6054600

El-Ashmawy, K. L. A. (2016). INVESTIGATION OF THE ACCURACY OF GOOGLE EARTH ELEVATION DATA. *ARTIFICIAL SATELLITES*, *51*(3). https://doi.org/10.1515/arsa-2016-0008

Hossain, F. (2018). Digital Elevation Modeling of Saint Martin Island, Bangladesh: a Method Based on Open Source Google Earth Data. *International Journal of Advanced Research*, *6*(2), 379–389. https://doi.org/10.21474/IJAR01/6449

Papasaika-Hanusch, C. (2012). *FUSION OF DIGITAL ELEVATION MODELS*. ETH ZURICH.

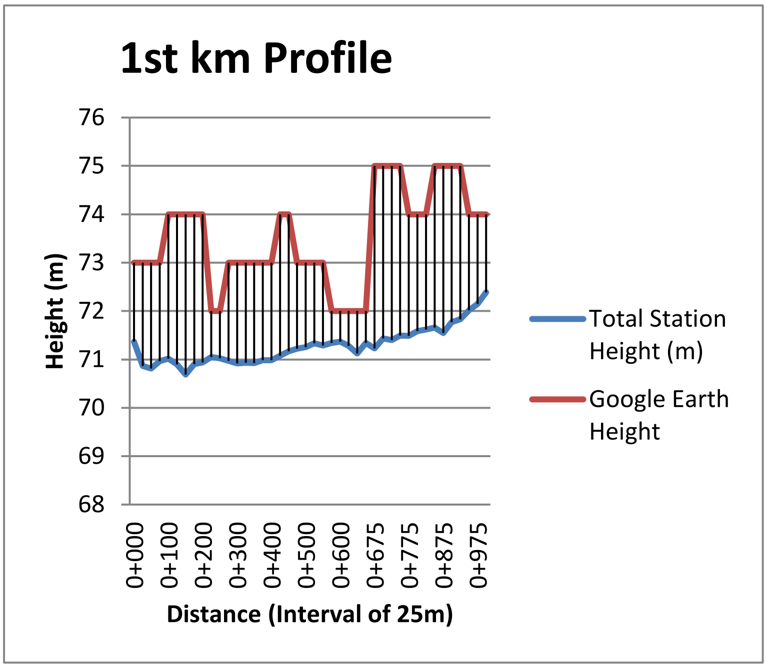
Richard, J., & Ogba, C. (2017). Analysis of Accuracy of Differential Global Positioning System (DGPS) and Google Earth Digital Terrain Model (DTM) Data using Geographic Information System Techniques Analysis of Accuracy of Differential Global Positioning System (DGPS) and Google Earth D. In *FIG Working Week 2017; Surveying the world of tomorrow - From digitalisation to augmented reality* (pp. 1–13).

Santillan, J. R., & Makinano-Santillan, M. (2016). Vertical accuracy assessment of 30-M resolution ALOS, ASTER, and SRTM global DEMS over Northeastern Mindanao, Philippines. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *41*(June), 149–156. https://doi.org/10.5194/isprsarchives-XLI-B4-149-2016

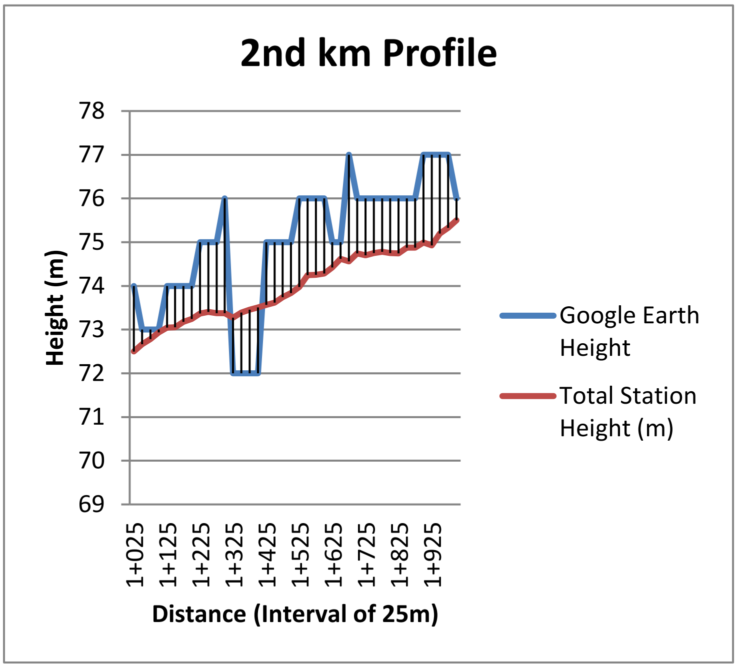
Wang, Y., Zou, Y., Henrickson, K., Wang, Y., Tang, J., & Park, B.-J. (2017). Google Earth elevation data extraction and accuracy assessment for transportation applications. *PLoS*

ONE, *12*(4), 1–17.
Wessel, B., Huber, M., Wohlfart, C., Marschalk, U., Kosmann, D., & Roth, A. (2018). Accuracy assessment of the global TanDEM-X Digital Elevation Model with GPS data. *ISPRS Journal of Photogrammetry and Remote Sensing, 139,* 171–182. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2018.02.017

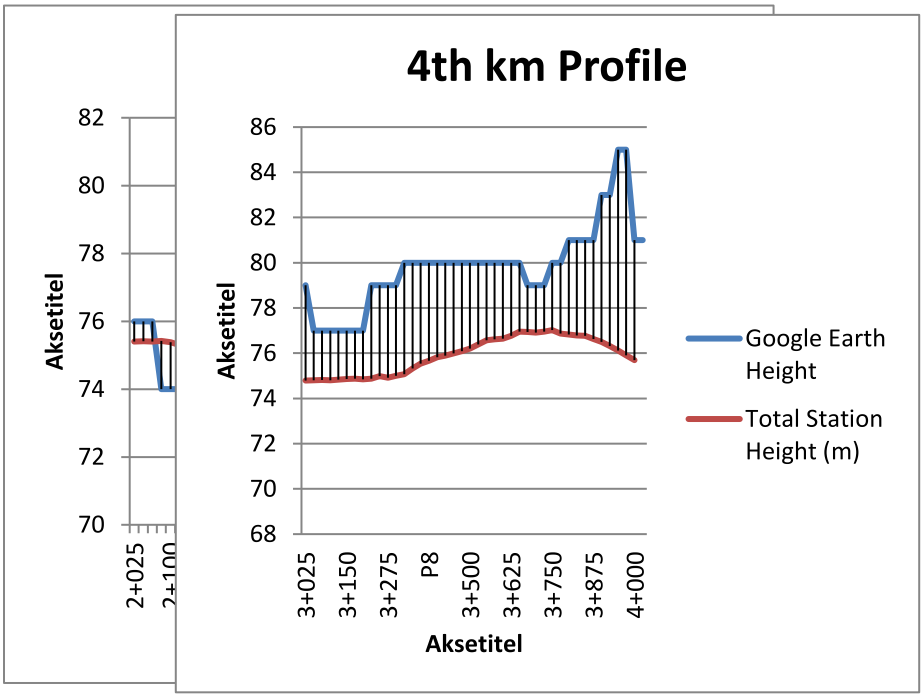
APPENDIX



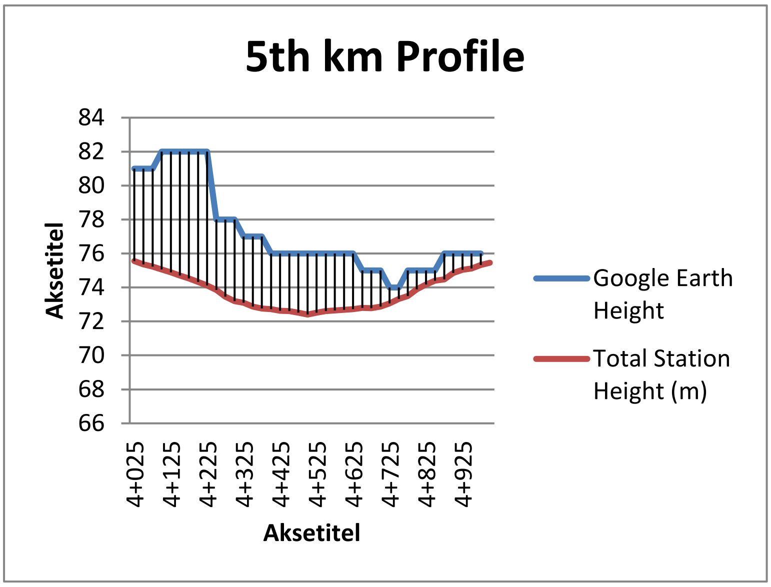
1st km Profile
Height (m)
76
75
74
73
72
71
70
69
68
0+000
0+100
0+200
0+300
0+400
0+500
0+600
0+675
0+775
0+875
0+975
Distance (Interval of 25m)
Total Station Height (m)
Google Earth Height

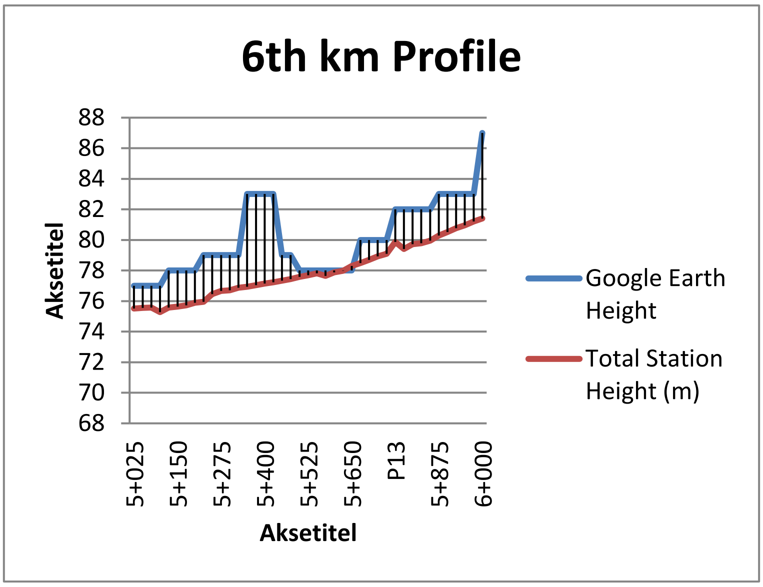


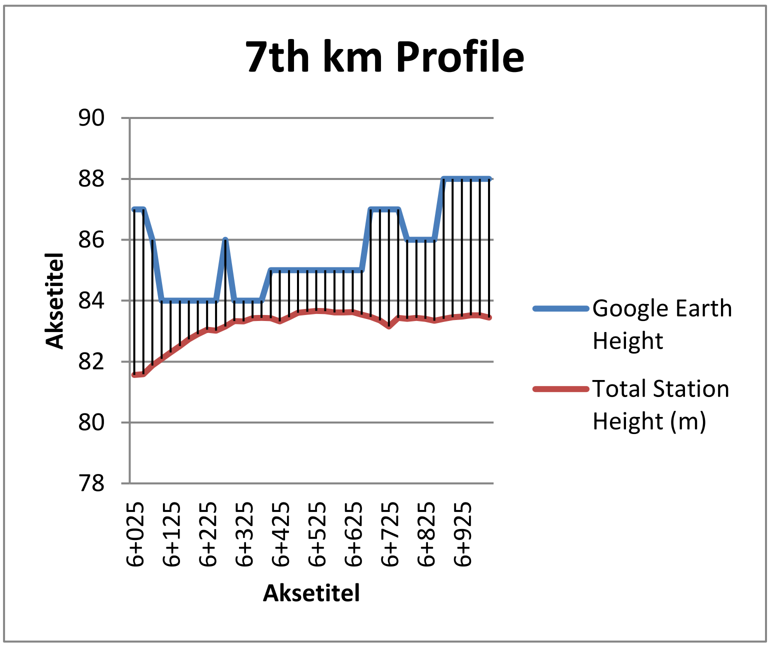
2nd km Profile
Height (m)
78
77
76
75
74
73
72
71
70
69
1+025
1+125
1+225
1+325
1+425
1+525
1+625
1+725
1+825
1+925
Distance (Interval of 25m)
Google Earth Height
Total Station Height (m)

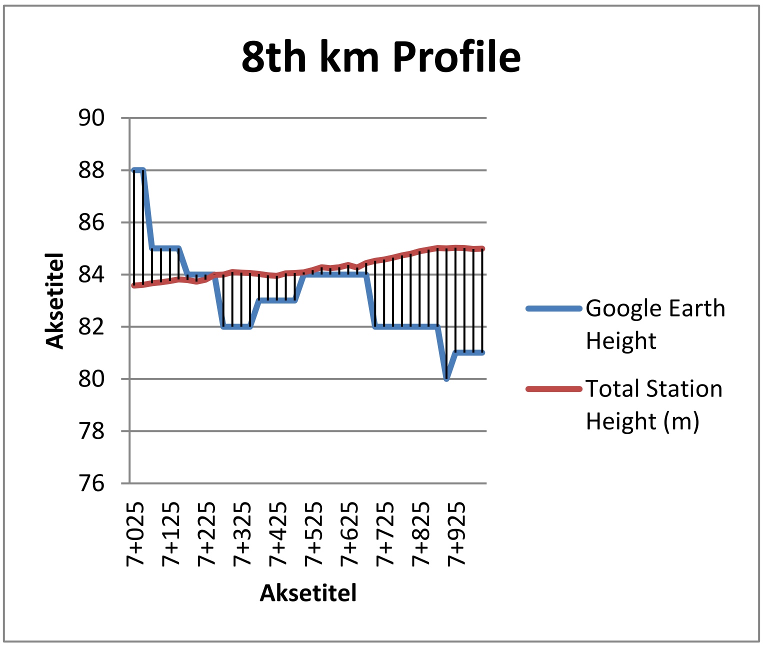


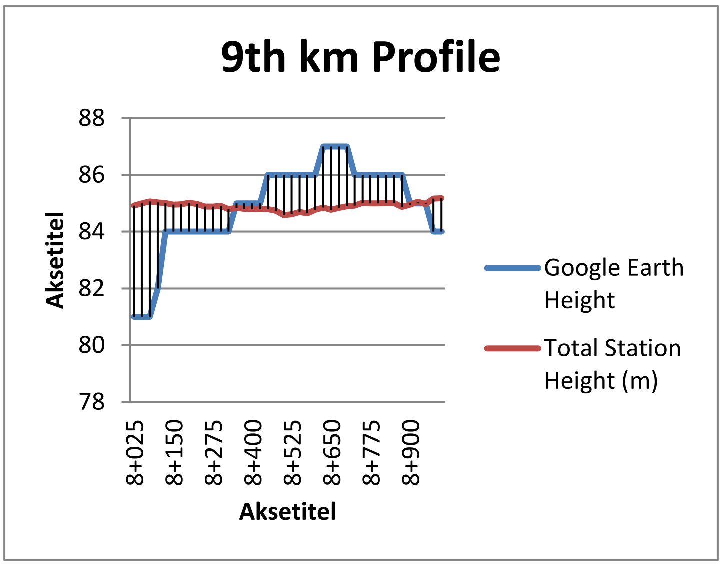
Aksetitel
82
80
78
76
74
72
70
2+025
4th km Profile
Aksetitel
86
84
82
80
78
76
74
72
70
68
3+025
3+150
3+275
P8
3+500
3+625
3+750
3+875
4+000
Aksetitel
Google Earth Height
Total Station Height (m)

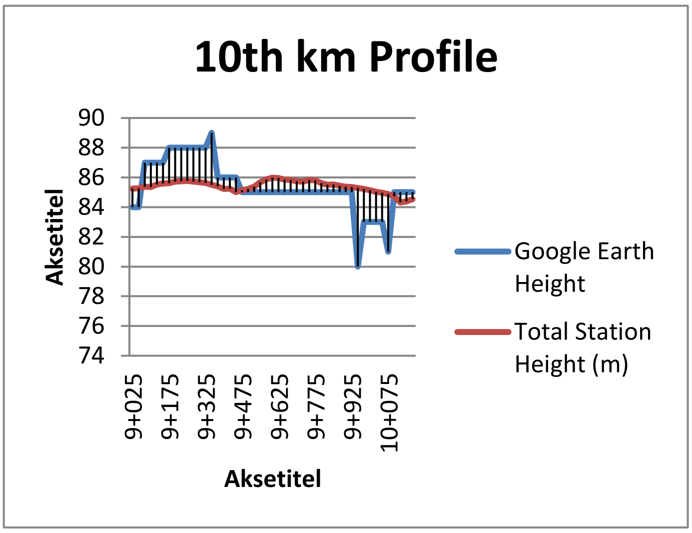

5th km Profile
84
82
80
78
76
74
72
70
68
66
Aksetitel
4+025
4+125
4+225
4+325
4+425
4+525
4+625
4+725
4+825
4+925
Aksetitel
Google Earth Height
Total Station Height (m)


6th km Profile
88
86
84
82
80
78
76
74
72
70
68
Aksetitel
5+025
5+150
5+275
5+400
5+525
5+650
P13
5+875
6+000
Aksetitel
Google Earth Height
Total Station Height (m)


7th km Profile
90
88
86
84
82
80
78
Aksetitel
6+025
6+125
6+225
6+325
6+425
6+525
6+625
6+725
6+825
6+925
Aksetitel
Google Earth Height
Total Station Height (m)


8th km Profile
90
88
86
84
82
80
78
76
Aksetitel
7+025
7+125
7+225
7+325
7+425
7+525
7+625
7+725
7+825
7+925
Aksetitel
Google Earth Height
Total Station Height (m)





CONTACT

Njike CHIGBU
Department of Surveying and Geo-Informatics
Abia State Polytechnic, Aba.
Abia State, Nigeria.
Department of Surveying and Geo-Informatics
Abia State University, Uturu.
Abia State, Nigeria.
Tel: +2348033423624
Email:njikec@gmail.com

Maduabughichi OKEZIE
Department of Surveying and Geo-Informatics
Abia State Polytechnic, Aba.
Abia State, Nigeria.
Tel: +2347030522657
Email:maduokezie7@gmail.com

Ikenna Donald ARUNGWA
Department of Surveying and Geo-Informatics
Abia State University, Uturu.
Abia State, Nigeria.
Tel: +2348035582298
Email: arungwaikenna@gmail.com
Email: ikenna.arungwa@abiastateuniversity.edu.ng