

INDIGENOUS COMMUNITY TREE INVENTORY: ASSESSMENT OF DATA QUALITY

M.F.Fauzi^a, N.H.Idris^{b*}, Din, A.H.M.^b, M.J.Osman^a, N.H.Idris^b, M.H.I.Ishak^c

^a TropicalMap Research Group, Faculty of Geoinformation and Real Estate, Universiti Teknologi Malaysia,
UTM Johor Bahru 81310 Johor

mfarid85@live.utm.my; mohamadjahidiosman@gmail.com

^b Department of Geoinformation, Faculty of Geoinformation and Real Estate, Universiti Teknologi Malaysia,
UTM Johor Bahru 81310 Johor - hawani@utm.my*, amihassan@utm.my, nurulhazrina@utm.my

^c Department of Control and Mechatronics, Faculty of Electrical Engineering, Universiti Teknologi Malaysia,
UTM Johor Bahru 81310 Johor - hafis@utm.my

KEY WORDS: Citizen Science, Tree inventory, Spatial Data Quality, Indigenous People, Volunteered Geographic Information

ABSTRACT:

The citizen science program to supplement authoritative data in tree inventory has been well implemented in various countries. However, there is a lack of study that assesses correctness and accuracy of tree data supplied by citizens. This paper addresses the issue of tree data quality supplied by semi-literate indigenous group. The aim of this paper is to assess the correctness of attributes (tree species name, height and diameter at breast height) and the accuracy of tree horizontal positioning data supplied by indigenous people. The accuracy of the tree horizontal position recorded by GNSS-enable smart phone was found to have a RMSE value of ± 8 m which is not suitable to accurately locate individual tree position in tropical rainforest such as the Royal Belum State Park. Consequently, the tree species names contributed by indigenous people were only 20 to 30 percent correct as compared with the reference data. However, the combination of indigenous respondents comprising of different ages, experience and knowledge working in a group influence less attribute error in data entry and increase the use of free text rather than audio methods. The indigenous community has a big potential to engage with scientific study due to their local knowledge with the research area, however intensive training must be given to empower their skills and several challenges need to be addressed.

1. INTRODUCTION

Tropical forest is one of the most complex ecosystems exist on Earth and has become a habitat of more than 50% of known species of flora and fauna. However, this forest is now threatened by extinction, degradation and deforestation due to human exploitation and disease (O'Hare *et al.*, 2014). Periodic observation and data collection have been conducted at certain area to monitor the rate of tree survival and growth.

The problem faced by organisation and authorities in gathering sufficient data for analysis is the data. Periodically data gathering and observation utilizes time and money (Halopainen *et al.*, 2014; Holmström *et al.*, 2003; Karila *et al.*, 2015; Pouliot *et al.*, 2002). Nowadays, many researchers have implemented the usage of remote sensing technique such as airborne LiDAR, unmanned aerial vehicle (UAV), terrestrial laser scanning (TLS) and optical satellite data (Bauwens *et al.*, 2016; Karila *et al.*, 2015; Liu *et al.*, 2014; Rahfl *et al.*, 2014; Solberg *et al.*, 2015). However, these techniques relies on estimation and calculation (Liang *et al.*, 2012), hence ground truth data collection is still required in checking and producing correct and valid data.

Citizen science has been used as a source of data gathering to supplement authoritative data. Citizen science is an initiative to engage community via data collection and observation. The programs tend to have certain objectives associate with

scientific studies, campaign and environmental monitoring. One good example of this initiative is the annual 'Christmas Bird Count' event where citizens were trained to identify and collect the attributes of bird species within their community to assist in a scientific study of bird populations (LeBaron, 2007). Subsequently, this practise has been adopted in analysing a series of satellite images for detecting objects that might have been relevant to the missing MH370 Malaysia airplane (Fishwick, 2014);

Many studies have been devoted to engage literate citizen in data collection. Citizen science initiative in tree inventory has become increasingly common, such as in Foster and Dunham (2014), Fuhrman *et al.* (2014) and Abd-Elrahman *et al.* (2010). Haw (2014) has identified several organisations that launched citizen science projects such as 'Ancient Tree Hunt' and 'Nature's Calendar' led by the UK Woodland Trust, 'Plant Tracker' led by the University of Bristol and 'Recording Invasive Species Count' led by the National Biodiversity Network Biological Records Centre. New York City has organised the 'Young Street Tree Mortality' to gather data to examine the factors that influence the survival of planted street trees (NYC, 2010).

However, a few studies attempt to engage semi-literate citizen in data collection. For examples, Kayapo indigenous people in Amazonia, State of Para, Brazil, collect the coordinates of

* Corresponding author

locations where illegal practices occur, mapping the ‘bee keeper’ residences, water resources, vegetation species as well as contaminant elements in 11 Brazilians states (Brito et al., 2013); mapping the slums areas in Phnom Penh, Cambodia., Tanzania and Zimbabwe (Roaf et al., 2005); engaged the rural citizens and indigenous groups in participatory mapping activities in Peru, the Philippines, Indonesia, Thailand and Malaysia (Corbett, 2009).

Goodchild (2007) stated that citizen scientist requires fairly high level skill to identify and collect the tree data and to control the accuracy of contributed data. In tree inventory, it is important to train a citizen scientist to follow the established protocol (Vogt and Fischer, 2014). Nevertheless, the accuracy of data contributed by citizen scientist has been highlighted in the literature, such as by Crall et al. (2011) and Fuhrman et al. (2014).

The present research explores the potential of indigenous people to engage in a scientific research of tree species identification for carbon mapping. The Royal Belum State Park (RBSP) is a reserved forest where human activities and access are limited, has been selected as the case study. The indigenous people were trained as citizen scientists to assist in tree inventory. The question is up to what extent the quality of data can be contributed by this community?

This paper discusses the accuracy of tree inventory data including the tree location, diameter at breast height (DBH), height and tree name contributed by indigenous people using GeoTrees mobile application.

2. METHODOLOGY

2.1 Study Area and Respondents

This study was conducted at Sungai Papan, Royal Belum State Parks (RBSP), Perak, Malaysia (5° 38' 05" N, 101° 24' 06" E) between 2nd and 8th November, 2015. Royal Belum is one of the reserve forests in Malaysia which is managed by the Perak State Park. A 200 x 300 metres training site was used to establish quadrants of 20 x 20 meter and sub-quadrants of 10 x 10 meter sizes near Sungai Papan Base Camp. Figure 1 shows a part of the plots and the position of trees observed in this study. There were 19 tagged trees in the plot. Different trees were allocated for different tasks.

This study solicited six volunteered respondents (n=6) to participate in this study. They were indigenous people that live in Kg Desa Damai and Kg Desa Permai, Gerik Perak. The respondents were all male within 20 - 55 years old which comprised of three respondents (n=3) from youth group (aged 15 - 40 years old¹) and another three respondents (n=3) are from adult group (40 - 55 years old). Two respondents attended secondary school (till form 1). The remaining respondents only attended primary school where one respondent attended until standard 6. The other three respondents attended school until standard 1. The respondents have been trained to collect tree

¹ This Malaysian definition of youth is based on 1997 National Youth Development Policy (NYDP) (KBS, 1997). However, starting 2018 Malaysia via Ministry of Youth and Sports will implement new age limit for youth as Malaysian Youth Policy (MYP) will replace the NYDP. The MYP age limit is from 15 to 30 compared to 15 to 40 years old under the NYDP (KBS, 2015).

data using the GeoTrees Android mobile data collector, Nikon Forestry and diameter tape.

Sg Papan Plot for Tree Monitoring

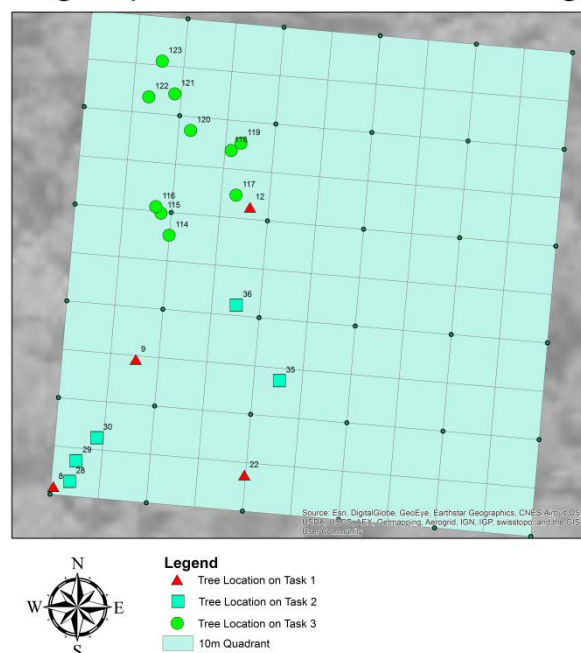


Figure 1. The plot used in this study

2.2 Research Framework

The respondents recorded the trees data, particularly the tree name, diameter at breast height (DBH) and height using the GeoTrees mobile application. The tree positioning was recorded automatically using location based services (LBS) on the smartphone when they keyed-in and submitted the data into the local server.

This study consists of three tasks. The first task required respondents to individually insert four flagged tree data (i.e. species name, DBH and height) into the mobile GeoTrees application. In Task 1, data were collected by researcher. The respondents are required to copy the tree data provided in the tree inventory worksheet. The second task required the respondents to collect the trees data individually. In this task, five trees were tagged by the respondents and the information was recorded individually. The third task required respondents to work in group and collect the tree data from ten tagged trees. In this task, two groups were formed where the members will have at least one adult and one member that could identify tree species.

The purpose of task 1 was to assess whether indigenous people could insert data accurately into a mobile data collector (in a condition of accurate reference data (height, DBH, species name) were given). The task 2 was to assess the accuracy of data collected by indigenous people independently and individually (in a condition that each respondent need to identify the tree name, independently). The finding in task 2 led to the design of task 3 which was to assess the accuracy of data collected by a group of indigenous people (in a condition that at least one member has the skill in species identification).

This study relies on several sources to act as reference data for quality comparison. The reference data for horizontal tree

positioning comparison were obtained from traverse surveys conducted by the TropicalMap Research Group in the study area. The reference data of species name was obtained from the appointed arborist, Mr Abu Rosnizam from the Forest Research Institute of Malaysia (FRIM). The reference data of species name in Orang Asli dialects were obtained from the Tok Batin Kg Desa Damai, Mr Ibrahim bin Angah.

3. INSTRUMENTS

3.1 GeoTrees Mobile Application

The GeoTrees mobile application from the previous study (Fauzi et al., 2016) was used to collect and stored the tree inventory, as shown in Figure 2. Using this application, the respondents will insert tree inventory data including tree label, species name, height, diameter at breast height (DBH) and tree images. The tree location (in latitude and longitude) was automatically recorded when data was inserted into the local database. Figure 3 shows the framework of storing data storing using GeoTrees mobile application.

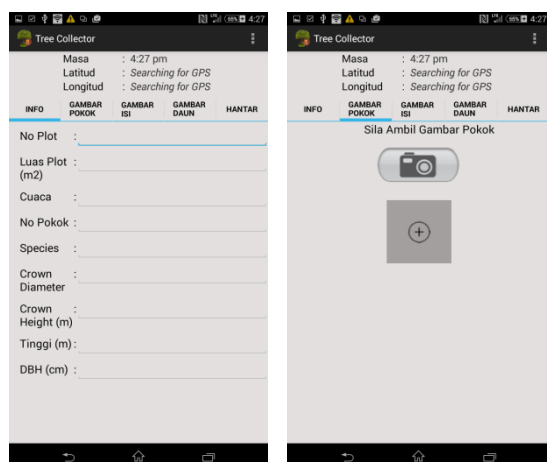


Figure 2. The User interface of GeoTrees mobile application



Figure 3. The workflow of data entry using GeoTrees mobile application

3.2 GPS Static Receiver

The ground control points (GCPs) established in the training site near Sungai Papan Camp was collected using this instrument. The values were obtained using static GPS. Two static GPS receivers were used where one static GPS receiver station was set up at a new control point while another receiver was stationed at the benchmark. The latter station acts as the base station for one hour to obtain the accurate positioning data. The GCPs were then used to locate the details of plot.

3.3 Total Station ES Series

The Total Station ES Series was used to collect the details of the plot and to obtain the location of tagged trees. The positioning tree points were transferred from GCPs.

3.4 Samsung Galaxy J

This study used a low cost GNSS-enabled smartphone, Samsung Galaxy J. It has a GNSS-enabled chipset that can receive signal from GPS and GLONASS. This chipset uses L1 frequency and C/A code to obtain the location. In this study, the device was installed with GeoTrees application as trees data collector.

3.5 Nikon Forestry 550

Nikon Forestry 550 was designed mainly for forestry use to meet the demand for angle compensated distance measurement. The built-in inclinometer provides easy readings of a tree-height, vertical separation (the difference in height between two targets), horizontal distance and angle in addition to actual distance. Measurement results are displayed on both internal and external LCD panels (Nikon, n.d.).

3.6 Diameter Tape (D-Tape/DBH Tape)

The diameter measuring tape enables a quick and easy calculation of tree diameters. It is used to measure cylindrical objects such as pipe and tree trunk. This instrument assumed the cylinder object was a perfect circle. The diameter tape provides an approximation of diameter.

To measure the diameter of a tree, the diameter tape (diameter side facing user) was wrapped around the tree, in the plane perpendicular to the axis of the trunk at 1.4 m above ground. Depending on the position of where the number "0" aligns with the rest of the tape, the diameter can be read directly from the tape (Kuers et al., 2012).

4. RESULTS AND DISCUSSIONS

4.1 The Comparison of Overall Errors of Tree Inventory according to Specific Task

4.1.1 Assessment of Task 1: In task 1, supervised individual tree data collection was conducted. The purpose of task 1 was to assess whether indigenous people could insert data accurately into a mobile data collector (in a condition of accurate reference data (height, DBH, species name) were given). The root mean square errors (RMSE) of tree positions recorded by the smartphone were calculated against the reference positioning data. All tree positioning data were obtained at the same time period maintaining the same scenario and condition. The errors of tree DBH and height were compared against reference data (the values have been prepared by the researcher before the assessment).

Data Entry	Overall RMSE
Tree location Easting	7.051 m
Tree location Northing	3.957 m
Average positioning error	8.085 m
DBH	1.639 cm
Tree Height	3.041 m

Table 1. Overall RMSE in tree inventory of task 1

As depicted in Table 1, the overall RMSE value was determined by averaging the RMSE values for each data entry of tree location, DBH, and tree height of tropical forest trees. The overall RMSE for tree DBH has lowest value compared with other attributes, while the tree position indicated ± 8 m of overall RMSE. Fauzi et al. (2016) have discussed several factors that might influence the larger error of tree position when using GNSS-enable mobile devices such as the issues of satellite geometry and diameter of tree canopy in the forests.

4.1.2 Assessment of Task 2: In task 2, unsupervised individual tree data collection approach was applied. In this task, all respondents were given a set of five tagged trees and were asked to identify tree name. The purpose of this task was to assess the skill of indigenous people in identifying tree names either at genus or species levels. As shown in Table 2, the overall RMSE of tree position in task 2 generate ± 8 m error as in task 1.

Data Entry	Overall RMSE
Tree location Easting	6.372 m
Tree location Northing	5.719 m
Average positioning error	8.562 m

Table 2. Overall RMSE of data entry in task 2

Respondent	Correct	Incorrect	Not Answered	Unknown
1	1	3	0	1
2	2	1	2	0
3	1	2	1	1
4	0	0	5	0
5	2	0	3	0
6	2	0	3	0

Table 3. The frequencies of 'not answered', 'correct' and 'incorrect', 'unknown' answer of tree names in task 2

Table 3 shows the number of correct and incorrect tree names given by each respondent. From the demographic data in Table 4, the respondent that provide the most correct answer are respondent 2, 5 and 6. Respondent 5 and 6 are adults with ages of 45 and 50 respectively and they attended school only for 1 year. However, respondent 2 is young but has attended school for 5 years and has some experience working as part time in logging industry.

4.1.3 Assessment of Task 3: Previous tasks have identified the limitations of each respondent. The ability of handling the application, collecting physical data and knowledge on species name are different between respondents. Therefore, in the task 3, data collection was conducted by group. Each group has three members which was given task to measure tree DBH, tree and species identification.

In group 1, respondent 1 was able to handle the mobile data entry and measuring the tree DBH. He used to do part time work in logging activity. Respondent 2 measured the tree height as he could handle the Nikon Forestry 550 very well in the previous task. Respondent 3 has a skill in tree identification in this group.

Meanwhile in group 2, respondent 6 has some knowledge in tree identification and respondent 5 could handle the Nikon Forestry 550 to measure the tree height. For DBH measurement, the respondent 4 has some experience working with tree measurement at another project in Tasik Kenyir. Table 4 below shows the demographic of the respondents.

Group	Respondent Id	Age	Education	Phone
1	1	22	Form 1	No
	2	27	Standard 5	Yes
	3	55	Standard 1	No
2	4	21	Form 1	Yes
	5	45	Standard 1	Yes
	6	50	Standard 1	Yes

Table 4. Demographic of respondents by groups

As in Table 4, the respondents 1 and 4 that measured tree DBH and able to make data entry via a smartphone have higher education background until form 1. They led the groups, able to read and write and could handle the smartphone quite well. Most respondents have at least a traditional phone. The elder respondents (i.e. respondent 3 and 6) identified tree names as they have some knowledge regarding tree identification. The other respondents cannot identify most of the tree names. Table 5 below shows the overall errors in task 3.

Data Entry	Overall RMSE
Tree location Easting	6.158 m
Tree location Northing	5.465 m
Average positioning error	8.234 m
DBH	12.214 cm
Tree Height	18.075 m

Table 5. Overall RMSE of data entry in task 3

As shown in Table 5, the overall RMSE of tree positions in task 3 generate $\pm 8m$ error as in task 1 and task 2. The errors of tree DBH and height were both higher. This study can conclude that that the 19 tree positions recorded in the Royal Belum Reserve Forest produced $\pm 8m$ error against reference data.

The positioning errors were consistent in the three tasks. This finding in line with Fauzi et al. (2016) which argued the horizontal positioning accuracy of using GNSS-integrated smartphone is mostly accurate within 5 to 15m in locating tree position. Several factors such as the satellite geometry signal shading by larger and dense tree canopy might introduce more errors to the recorded positions (Yahya and Kamarudin, 2008).

4.2 The Accuracy of Tree Names

One important parameter that commonly used in tree inventory is a tree name either genus (common name) or species levels. Table 6 below shows the result of tree names between reference and collected data.

	Correct	Incorrect	Not Answered	Unknown
Task 1	14	1	1	0
Task 2	8	6	14	2
Task 3	4	5	0	3

Table 6. Comparison analysis between reference and collected data of tree names

From all three tasks in recording 58 tree names, about 44.8% tree names collected by indigenous people were correct while another 20.7% were incorrect. About 8.6% of the names flagged as unknown as this study could not identified the correctness of the data. This problem occurred due to some respondents given the tree names using their local dialects which were either Jahai or Temiar.

The individual task 2 has shown the highest number of respondents as not able to supply tree names. This indicates their lack of knowledge in tree species identification. Through limited sample, this study found that the younger respondents have a problem with trees identification. According to Crall et al (2011, p.439), 'accurate taxonomic identification requires years of specialized training and remains a barrier of data quality among diverse collector'.

To overcome this problem, the respondents were divided into groups where each group combines youth and adult respondents. As shown in Table 6, the number of 'not answered' has been reduced dramatically where both groups provide the tree names mostly in their dialects. Table 7 shows the method used by respondents/groups to store the tree species name into the mobile application.

	Free text	Audio	Not Answered
Task 1	15	0	1
Task 2	8	12	10
Task 3	12	0	0

Table 7. Data entry method used by respondents.

Tree No	Group 1			Group 2			Tree Name Reference (Arborist)
	Tree Name supplied by the Indigenous people	Tree Name (Malay)	Correctness	Tree Name supplied by the Indigenous people	Tree Name (Malay)	Correctness	
1	<i>Kep1</i>	<i>Rambutan Hutan</i>	✗	<i>Kif</i>	<i>Rambutan Hutan</i>	✗	<i>Redan</i>
2	<i>Gong</i>	--Unknown--	✗	<i>Lalok</i>	<i>Mersawa</i>	✗	<i>Kenanga Hutan</i>
3	<i>Jangkang</i>	<i>Simpur Jangkang</i>	✗	<i>Jankap</i>	<i>Simpur Jangkang</i>	✗	<i>Mentulang</i>
4	<i>Gong</i>	--Unknown--	✗	<i>Kupang</i>	<i>Kedondong</i>	✓	<i>Kedondong Gergaji</i>
5	<i>Telas</i>	<i>Keranji</i>	✓	<i>Telas</i>	<i>Keranji</i>	✓	<i>Keranji</i>
6	<i>Tangbai</i>	<i>Mempisang</i>	✓	<i>Tambai</i>	<i>Mempisang</i>	✓	<i>Mempisang</i>
7	<i>Gong</i>	--Unknown--	✗	<i>Tabung</i>	--Unknown--	✗	<i>Sengkuang</i>
8	<i>Pal</i>	--Unknown--	✗	<i>Katong</i>	<i>Kekatung (Keranji)</i>	✗	<i>Mempisang</i>
9	<i>Kenyar Ikek</i>	--Unknown--	✗	<i>Perahot</i>	--Unknown--	✗	<i>Sebasah</i>
10	<i>Tengkel</i>	<i>Penarahan</i>	✗	<i>Tonkul</i>	<i>Penarahan</i>	✗	<i>Mempisang</i>
		<i>Correct</i>	2		<i>Correct</i>	3	
		<i>Incorrect/Unknown</i>	8		<i>Incorrect/Unknown</i>	7	

Table 8. Tree names provided by groups in task 3 and the reference data

From Table 7, the use of free text data entry to contribute the data has significantly increased in task 3. The number of correct tree names has increased (as shown in Table 7) from 20% in task 2 to 33.3% in task 3 while number of tree names not answered has reduced to zero.

Based on Table 8, group 1 provided two correct tree names (against the reference data supplied by the expert) while group 2 has three correct tree names respectively. The number is low that represent only 20% to 30% correctness. While the rest are either incorrect tree names or cannot be identified. Table 8 shows the comparison of the tree names between groups and against the tree names identified by the expert (arborist). From the table, 40% of data provided between groups are similar (but have slight differences in the spelling), 10% of data is accurately matched and another 50% are not similar. There are also tree names that were similar provided by both groups but did not match with the reference data (identified by the arborist) such as tree labelled 1 (*Rambutan Hutani*), 3 (*Simpur Jangkang*) and 10 (*Penarahan*) as shown in Table 8.

4.3 The Accuracy of Tree Height

Another attribute collected in tree inventory is trees height. In this study, only task 1 and task 3 involved recording the trees height. In task 1, the respondent's ability to operate and store the collected data was observed. Table 9 shows the individual data stored by respondents and the associate errors.

No	Tree Height		Remark
	Indigenous People	Reference	
1	13.o m	13.0 m	Typing error (mistaken 0 with 'o')
2	13 .0 m	13.0 m	Typing error (space between 3 and decimal point)
3	130m	13.0 m	Typing error (no decimal point)
4	13òm	13.0 m	Typing error (mistaken 0 with 'o') and no decimal point)
5	9.8m86145 cm145cm	9.8 m	Typing error (mixed with other data)
6	242m19.5c m	24.2 m	Typing error (mixed with other data)

Table 9. A list of errors detected in data entry of tree height (in task 1) against the reference data.

In this task, respondents were required to copy tree height data from a paper based worksheet into the digital database via mobile application. In Table 9, the identified errors were associated with spacing and decimal point issues and confusion between letter '0' and number 'zero'. Several reasons might influence to these errors such as their familiarity of data entry using smartphone and the education background. However, this was their first task of data entry using mobile device. A few trials could be given to avoid this error. This can be seen when no such error emerged in task 3. The 10-minutes of data entry training before the assessment conducted is not sufficient and further training could be conducted to avoid these errors.

Based on the result in Table 10, it shows that group 1 provide better result with RMS error of 2.233m than group 2 which

produced RMS error up to 25.440m. The values given by Group 1 were slightly underestimate while group 2 provided highly overestimate reading. As shown in Figure 4, the tree labelled 1 until 4 indicate low biases which were 1.8m, -1m, 1.4m and 0.8m. However, for tree labelled 5 and 6, there were very large biases which then contribute to high RMS error.

	Bias	RMS error
Group 1	0.683	2.433
Group 2	-15.133	25.440

Table 10. Bias and RMS error according to group for tree height data in task 3.

Based on ground observation, these two trees were surrounded with many trees. This situation might not be a problem for experienced individual but for a citizen scientist who has just been exposed to measurement technique using Nikon Forestry 550, have the tendency to use different target tree was obstructed by other object. Hence, larger errors were produced.

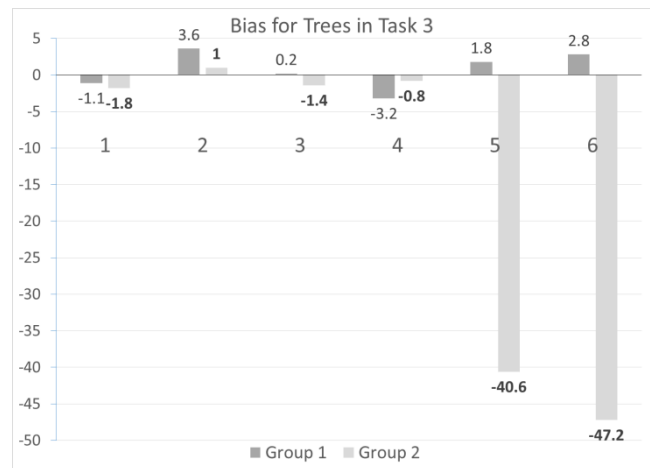


Figure 4. The bias of tree height in task 3.

As shown in Figure 5, it is crucial to maintain the same target when measuring the tree height using the Nikon Forestry as the measurement calculate the distance of the subject to the object (tree) at 1) eye level, 2) the top and 3) bottom of the tree.

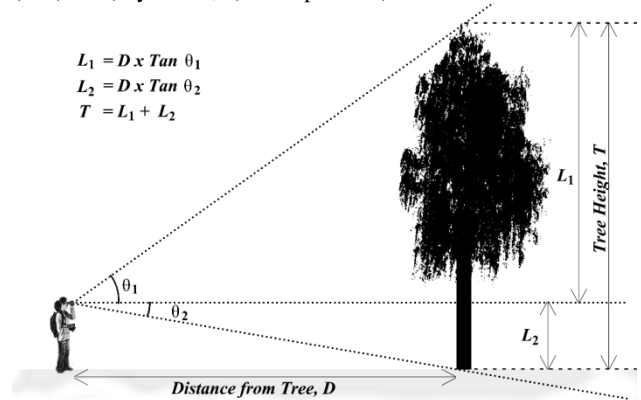


Figure 5. Tangent method that was used for calculating the tree height

4.4 Accuracy of Tree Diameter at Breast Height (DBH)

Tree DBH is another parameter commonly collected in tree inventory, particularly in a study to estimate carbon storage and biomass. In this study, only task 1 and task 3 involved recording

the tree DBH. Table 12 shows the list of tree DBH errors found in task 1.

No	Tree DBH		Remark
	Collected	Reference	
1	22.2 m	22.2 cm	Wrong unit
2	22.2 m	22.2 cm	Wrong unit
3	2232m	22.2 cm	Wrong unit and no decimal
4	2222m	22.2 cm	Wrong unit and no decimal
5	145cm	14.5 cm	No decimal point
6	12	14.5 cm	Incorrect value input
7	195cm	19.5 cm	No decimal point

Table 12. The list of errors associates to tree DBH measurement in task 1 and the reference data.

From Table 12, the respondents seem confused with the measurement unit for tree DBH whether in centimetre (cm) or meter (m) although the units were clearly stated in the worksheet. There was error associate with decimal point. However, as in tree height assessment, this error did not occur in task 3. As argued by Crall et al (2011), the range of experience and skill levels of contributors in citizen science program attributed to the level of accuracy and quality of data collected.

The respondents used DBH tape to measure tree DBH of the selected trees. Using DBH tape was not as complex as Nikon Forestry 550 as the respondents able to measure and collect diameter of all trees at the breast height. Table 13 shows the bias and RMS error of the recorded data.

	Bias	RMS error
Group 1	-0.200	4.845
Group 2	-6.450	16.580

Table 13. Bias and RMS errors according to group for tree DBH measurement in task 3.

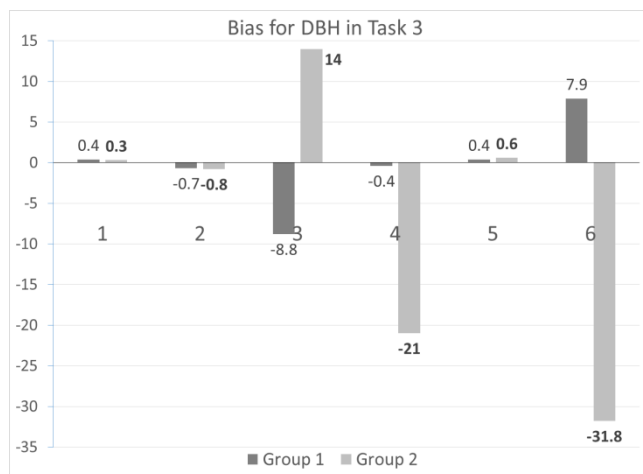


Figure 5. The bias of tree DBH values for all trees in task 3.

Even though measuring the tree DBH was very straight forward, some of measurements collected by respondents were not very

good. About 66.7% and 50% of the tree DBH data collected by group 1 and 2, respectively produced error less than 1cm. This represents 58.3% of all tree DBH measurement in task 3. This bias result is proportional with RMS error in Table 12 which shows that measurement made by group 1 produced the lowest error than group 2. However, data accuracy in group 2 was lower most likely because of a few values of tree DBH in the dataset (as in Table 13) and tree DBH measurement for tree labelled 3, 4 and 6 (as shown in Figure 5) were significantly different against the reference data.

5. CONCLUSION

This paper discusses the quality of tree inventory data contributed by semi-literate indigenous people for scientific research program. Through limited sampling, several findings associate with the accuracy of horizontal positioning accuracy of tree location (detected via GNSS integrated smartphone) and correctness of attribute (tree name, DBH and height) have been identified. Nevertheless, further study need to be conducted so that the sample size is significant to represent the population of indigenous people, particularly at the Royal Belum Reserve Forest. The tree position produced $\pm 8m$ errors; hence the smartphone is not suitable to accurately record individual tree position for data inventory, particularly in dense tropical forest. However, the supplementary data that were collected associated with the position of quadrant and sub-quadrant and the tree tag on a 10 x 10 m grid plot could use as a basis to estimate a tree position. The position recorded by a smartphone is only suitable for cohort study that monitors a group of trees.

Although limited sampling, the education background has not effect on their knowledge on species identification. However, the knowledge they gained based on their working experience can influence their skills in tree identification. This study shows that adult respondents who have been working for years have more knowledge in species identification than the younger respondents. However, younger respondents were able to handle the instruments to measure and record the tree DBH and height after short training were given. To conclude, indigenous people at Royal Belum have a big potential to act as citizen scientist in tree species identification, particularly due to their local knowledge within the research area. However, several challenges have been identified in this study that need to be tackled before the vision could be achieved.

ACKNOWLEDGEMENTS

This work was supported by the the Ministry of Higher Education Malaysia and the Universiti Teknologi Malaysia under UTM RG Flagship Grants [vote number 02G17], GUP Tier-1 [vote number 10H79] and FRGS Grant [number 4F656]. The authors would like to thank the Perbadanan Negeri Perak and Orang Asli community at Royal Belum Reserve Park to support this study.

REFERENCES

Abd-Elrahman, A. H., Thornhill, M. E., Andreu, M. G., and Escobedo, F., 2010. A community-based urban forest inventory using online mapping services and consumer-grade digital images. *Int J Appl Earth Obs Geoinf*, 12(4), pp. 249–260. <http://doi.org/10.1016/j.jag.2010.03.003>

- Bauwens, S., Bartholomeus, H., Calders, K., and Lejeune, P. 2016. Forest Inventory with Terrestrial LiDAR: A Comparison of Static and Hand-Held Mobile Laser Scanning. *Forests*, 7(6), pp. 127. <http://doi.org/10.3390/f7060127>
- Brito, L. P., Reis, D. N.C., Brito, R. L., Fernandes, V. de O., and Alixandrini, M. 2013. "Empowering Brazilian Unprivileged Communities with High Geotechnological Tools." Proceedings of the 26th International Cartographic Conference – From Pole to Pole, Dresden, August 25–30. http://icaci.org/files/documents/ICC_proceedings/ICC2013/_extendedAbstract/316_proceeding.pdf.
- Corbett, J. 2009. *Good Practices in Participatory Mapping*. Rome: International Fund for Agricultural Development (IFAD).
- Crall, A.W., Newman, G.J., Stohlgren, T.J., Holfelder, K.A., Graham, J. and Waller, D.M. 2011. Assessing citizen science data quality: an invasive species case study. *Conservation Letters*. 4. Pp.433-442.
- Fauzi, M. F., Idris, N. H., Yahya, M. H., Din, A. H. M., Idris N. H., Lau, A. M. S. and Ishak, M. H. I. 2016. Tropical Forest Tree Positioning Accuracy: A Comparison of Low Cost GNSS-Enabled Devices. *International Journal of Geoinformatics*. 12(2), pp. 59-66.
- Fishwick, C. 2014. "Tomnod – the online search party looking for Malaysian Airlines flight MH370", *The Guardian*. <http://www.theguardian.com/world/2014/mar/14/tomnod-online-search-malaysian-airlines-flight-mh370> (16 March 2014)
- Foster, A., & Dunham, I. M., 2014. Volunteered geographic information, urban forests, & environmental justice. *Computers, Environment and Urban Systems*. <http://doi.org/10.1016/j.compenvurbysys.2014.08.001>
- Fuhrman, N. E., Morgan, A. C., Copenheaver, C. A., Peterson, J. A., Newberry, M. G., Deloach, S. G., and Gevel, S. V. D. 2014. Repeated Monitoring of Forest Plots: Evaluating the Accuracy of Student Scientist Data. *NACTA Journal*, (June), pp. 95–101.
- Goodchild, M. F. 2007. Citizens as sensors: the world of volunteered geography. *GeoJournal*, 69(4), pp. 211–221. <http://doi.org/10.1007/s10708-007-9111-y>
- Holopainen, M., Vastaranta, M., and Hyypä, J. 2014. Outlook for the next generation's precision forestry in Finland. *Forests*, 5(7), 1682–1694. <http://doi.org/10.3390/f5071682>
- Haw, K., 2014. The rise of citizen science: How can community research help nature?. *ECOS: A Journal of Conservation*. 35(2), pp. 12-21.
- Holmström, H.; Kallur, H.; Ståhl, G., 2003. Cost-plus-loss analyses of forest inventory strategies based on kNN-assigned reference sample plot data. *Silva Fennica*, 37, pp. 381–398.
- KBS, 1997. "National Youth Development Policy", <http://www.kbs.gov.my/my/akta-pekeliiling/dasar.html?download=35:national-youth-development-policy> (1 July 2016)
- KBS (Kementerian Belia dan Sukan Malaysia), 2015. "Malaysian Youth Policy", <http://www.kbs.gov.my/en/malaysian-youth-policy.html?download=253:malaysian-youth-policy> (1 July 2016)
- Karila, K., Vastaranta, M., Karjalainen, M., and Kaasalainen, S., 2015. Tandem-X interferometry in the prediction of forest inventory attributes in managed boreal forests. *Remote Sensing of Environment*, 159, 259–268. <http://doi.org/10.1016/j.rse.2014.12.012>
- Kuers, K., Lindquist, E., Dosch, J., Shea, K., Machado, J. L., LoGiudice, K., Simmons, J., and Anderson, L., 2012. "Permanent Forest Plot Project (PFPP) Version 1.05 (Appendix VI)", http://erenweb.org/wp-content/uploads/2012/08/EREN-PFPP_Appendix-VI_TreeDBHGuide_30July2012.pdf (30 July 2016)
- Le Baron, G., 2007. Audubon's Christmas Bird Count to 21 St Century Citizen Science. Citizen Science Toolkit Conference
- Liang, X., Hyypä, J., Kaartinen, H., Holopainen, M., and Melkas, T., 2012. Detecting Changes in Forest Structure over Time with Bi-Temporal Terrestrial Laser Scanning Data. *ISPRS International Journal of Geo-Information*, 1(3), pp. 242–255. <http://doi.org/10.3390/ijgi1030242>
- Liu, J., Hyypä, J., Yu, X., Jaakkola, A., Liang, X., Kaartinen, H., Hyypä, H., 2016. Can global navigation satellite system signals reveal the ecological attributes of forests? *International Journal of Applied Earth Observation and Geoinformation*, 50, 74–79. <http://doi.org/http://dx.doi.org/10.1016/j.jag.2016.03.007>
- Nikon, n.d. "Instruction Manual for Forestry 550", Tokyo, Japan. <https://cdn-1.nikon-cdn.com/pdf/sportoptics/Forestry.pdf> (25 July 2016)
- NYC. 2010. New York City's Urban Street Tree Mortality Study: Site Assessment Tools Description. https://www.nycgovparks.org/sub_your_park/trees_greenstreets/images/how_to_assess_your_citys_st_survival_10-6-2010.pdf (2 August 2016).
- O'Hare, G., Sweeney, J., and Wilby, R., 2014. *Weather, Climate and Climate Change: Human Perspectives*. Taylor and Francis.
- Pouliot, D.A.; King, D.J.; Bell, F.W.; Pitt, D.G., 2002. Automated tree crown detection and delineation in high-resolution digital camera imagery of coniferous forest regeneration. *Remote Sens. Environ.*, 82, pp. 322–334.
- Rahlf, J., Breidenbach, J., Solberg, S., Næsset, E., and Astrup, R. 2014. Comparison of four types of 3D data for timber volume estimation. *Remote Sensing of Environment*, 155, pp. 325–333. <http://doi.org/10.1016/j.rse.2014.08.036>
- Roaf, V., A. Bouvier, B. Calaguas, A. Etherington, and G. Burrows. 2005. *Community Mapping: A Tool for Community Organising*. London: WaterAid Press.
- Solberg, S., Riegler, G., and Nonin, P., 2015. Estimating forest biomass from TerraSAR-X stripmap radargrammetry. *IEEE Transactions on Geoscience and Remote Sensing*, 53(1), pp. 154–161. <http://doi.org/10.1109/TGRS.2014.2319853>
- Vogt, J. M., and Fischer, B. C., 2014. A Protocol for Citizen Science Monitoring of Recently-Planted Urban Trees A Protocol for Citizen Science Monitoring of Recently-Planted Urban. *Cities and the Environment*, 7(2).
- Yahya, M. H. and Kamarudin, M. N., 2008, Analysis of GPS Visibility and Satellite-Receiver Geometry over Different Latitudinal Regions, Presented at International Symposium on Geoinformation (ISG 2008), Kuala Lumpur, Malaysia.