USE OF OPEN DATA KIT AND ANDROID DEVICES FOR FAST AND EASY DATA COLLECTION – A CASE STUDY FROM THE PHILIPPINES

Ubo Pakes (1,3), Rhea Subong (2), Mary Ann Gumban (2), Klaus Schmitt (1), Mari Trix Estomata (1)

[1] Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, 9th Floor, Bank of Makati Building, Ayala Avenue Extension near corner Metropolitan Avenue, 1209 Makati City, Philippines
[2] College of Management, University of the Philippines Iloilo City Campus, Iloilo City, Philippines
[3] Center for Environmental Informatics (CENVI) University of the Philippines Cebu, Cebu City, Philippines
Email: ubo@gis-tm.com; rjsubong@up.edu.ph; mtgumban@gmail.com; klaus.schmitt@giz.de; mari.estomata@giz.de

KEY WORDS: ODK, Android, digital survey forms, spatial data collection

ABSTRACT Collection of spatial data in the field, data analysis, and mapping using Geographic Information Systems (GIS) is complex, time-consuming and expensive. The growing popularity of Android devices with GPS (Geographic Positioning System) and wireless internet (WiFi) capability, in combination with open-source web-based tools, has changed this fundamentally. Open Data Kit (ODK), a free, open source suite of tools for collection and analysis of spatial data was tested during a survey of cacao farms on Panay Island in the Philippines. The software enables users to easily design digital survey forms and collect data in the field using Android devices. Photos, videos or speech can also be recorded as part of a survey and the location of the data collection is recorded using the GPS of the Android device. Data entered in the survey form are automatically saved at the end of the survey and uploaded to the server database. If there is no network/internet connection, data can be uploaded later when users have connection to the internet. After uploading data to the server, they can be presented and analysed using specific features of ODK or visualised in Google Earth. Data can be exported for additional analysis and GIS mapping in other software. Some results from the cacao farm survey will be presented to illustrate this. Programming skills are not necessary to start an ODK project. Only a Google account and a free copy of JAVA are required to set-up an ODK server, however for data security reasons ODK should be used with a dedicated server and data encryption. ODK makes data collection, analysis and mapping faster and more efficient. This solves the bottleneck of many surveys, the time-consuming and error prone manual data entry, through automatic data upload to a server database. ODK is therefore suitable for fast and accurate field truthing of remote sensing projects and accuracy assessments of remotely sensed and analysed maps. As ODK is an open source software, it is constantly further developed and thus will continue to meet the needs of users in a fast-changing digital world.

1. INTRODUCTION

The collection of spatial data in the field changed fundamentally when handheld GPS (Geographic Positioning System) receivers became available for commercial use in 1989. During the 1990s, the GPS signal accuracy for civilian use was intentionally degraded through a feature called "Selective Availability". In May 2000, the Clinton administration ended the use of "Selective Availability", thus providing full commercial access to the US GPS satellite system with higher accuracy (Tomkiewicz et al. 2010, https://www.gps.gov/systems/gps/performance/accuracy/).

The accuracy commitment of the US government refers to the signals transmitted in space which currently has a global average user range error (URE) of ≤7.8 m, with 95% probability. However, actual performance of GPS devices exceeds this specification. On 11 May 2016, the global average URE was ≤0.715 m, 95% of the time (https://www.gps.gov/systems/gps/performance/accuracy/). Differential GPS devices and dual-frequency receivers (the receiver tracks more than one radio signal from each satellite on different frequencies) increase accuracy and can enable real-time positioning within a few centimetres, and long-term measurements at the millimetre level (El-naggar, 2011).

The growing popularity of Android devices with GPS and wireless internet (WiFi) capability, in combination with open-source web-based tools, has again fundamentally changed collection of spatial data in the field. GPS-enabled smartphones are readily available and typically accurate to within a 4.9 m radius under open sky (van Diggelen and Enge, 2015). This has been improved through the use of multiple Global Navigation Satellite Systems beyond US GPS such as Galileo (European Union), BeiDou (China), the Russian GLONASS (GLObal NAvigation Satellite System) and dual-frequency receivers which are available since 2018 and supported since the release of Android Nougat 7.0 in 2016 (Chen et al., 2019). Android devices are now able to perform surveys with decimetre-level accuracy and are therefore suitable for applications such as map updating and cadastral survey (Realini et al. 2017; Chen et al., 2019; Wu et al., 2019).

In addition, the use of Android devices in combination with open-source web-based tools has made data analysis, and mapping less complex, time-consuming and expensive than the use of Geographic Information Systems (GIS).

2. METHODS

The suitability of Open Data Kit (ODK, http://opendatakit.org), a free, open source suite of tools for collection and analysis of spatial data (Hartung et al., 2010; Brunette et al., 2013), was tested during a survey of cacao farms on Panay Island in the Philippines (Fig. 1). Five teams of two data collectors each collected data from 118 villages in 44 municipalities of the four provinces in Panay Island during a period of 21 days of field work in 2017. During this period, 252 farms were surveyed, and 197,718 cacao trees recorded (Pakes and Subong 2017; Pakes et al., 2018).

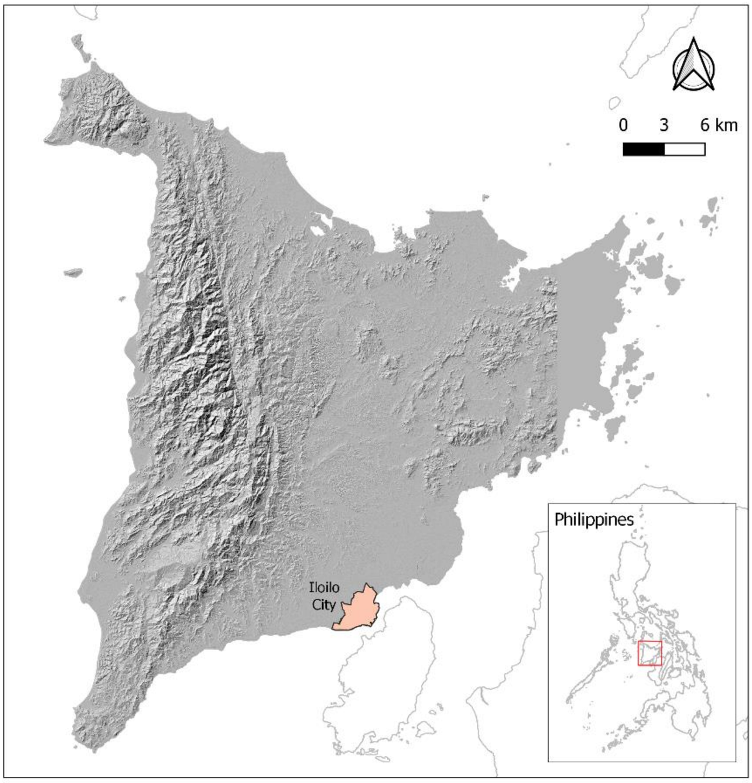

Figure 1, Panay Island and the Panay Mountain Range (West Side).

The survey forms were designed in Microsoft Excel (any other spread sheet application can also be used) and then rendered into a sequence of input prompts that apply form logic, entry constraints and repeating sub-structures using ODK Collect uses the OpenRosa standard for digital surveys. The forms were then uploaded to Android devices. During the data collection, users worked through the prompts (questions) while the location of the data collection is recorded automatically using the GPS of the device. Photos, videos or speech can also be recorded as part of a survey. The data entered in the form are automatically saved at the end of the survey and uploaded to the server database. If there was no network/internet connection, data were uploaded later as soon as the users had connection to the internet.

ODK Aggregate is the server that accepts the forms and the data collected. After uploading the data to the server, they were presented and analysed using specific features of ODK such as basic graphs and maps or visualised in Google Earth. ODK Aggregate was also used for data export through the download of datasets as comma-separated-value (CSV) files for additional data analysis and GIS mapping in other software.

ODK Briefcase was used for data backup from the web-based server to a computer. These data were then used for further analysis.

3. RESULTS

ODK makes spatial data collection, analysis and mapping faster and more efficient. Programming skills are not necessary to start an ODK project. Only a Google account and a free copy of JAVA are required to set-up an ODK server. However, for data security reasons, ODK should be used with a dedicated server and data encryption. This can ensure that confidential information and identities of individuals represented in data sets are protected from unauthorised access and manipulation by third parties.

Cacao farms and plantations are not evenly distributed throughout the Island. Figure 2 shows that most of the farms are located to the east of the Panay mountain range (see figure 1) and in relatively low-lying areas away from the coast. Generally, these farms are also bigger.

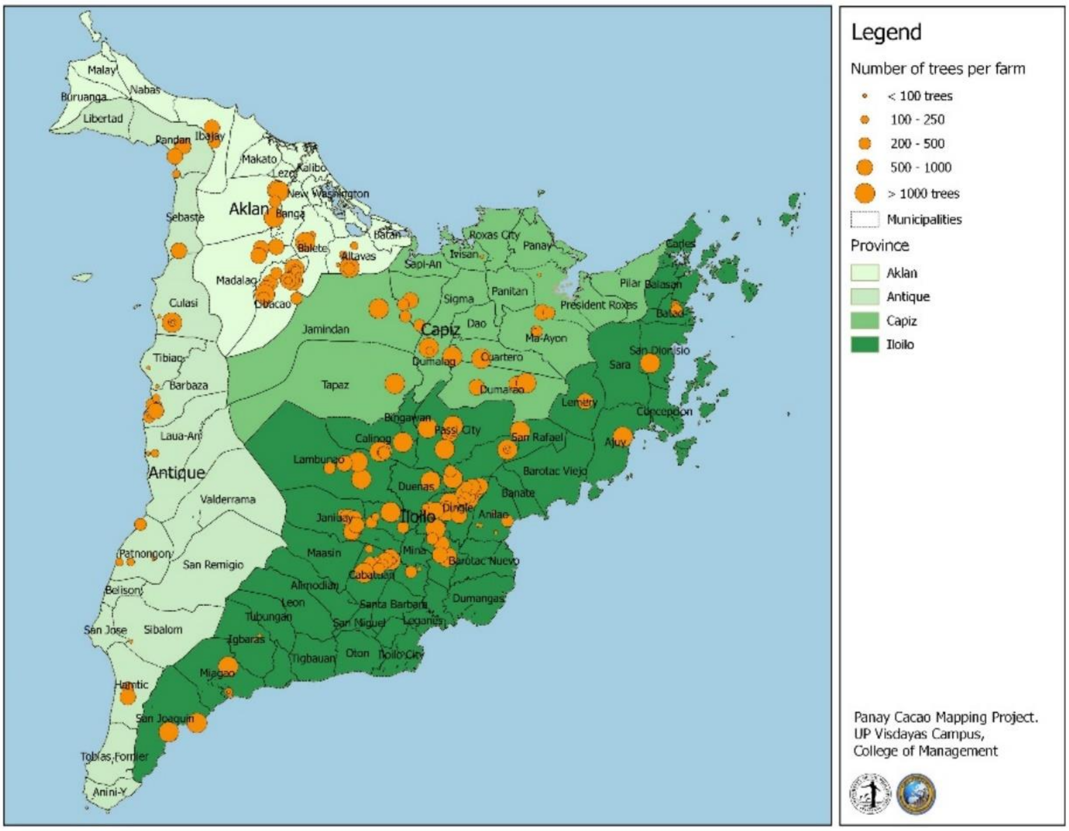

Figure 2, Location and Size of Cacao Farms in Panay Island.

Table 1 shows the years when the cacao trees were planted in Panay Island. The first cacao pods on a tree can be harvested 2 years after planting. Therefore, about 74,000 trees or about 37% of the capacity for cacao production in Panay was not yet utilised during the time of the survey because these trees were planted in 2016 and 2017.

Table 1, Planting Dates of the Majority of Cacao Trees in Panay.

Year planted	Number of trees	Percentage
before 2013	45,244	22.88
2013	20,783	10.51
2014	46,657	23.60
2015	10,606	5.36
2016	58,762	29.72
2017	15,666	7.92
Total	197,718	100

Cacao is a high value commodity in the global market (Department of Agriculture - Bureau of Plant Industry, 2017) and it is therefore important for local government units to consider its contribution to the local economy and its future potential. The mapping of cacao farms together with information about tree numbers and age brings insights for the local executives about the economic contribution of cacao to annual income and property valuation. This can guide local policy makers in developing intervention programmes that will help farmers and producers in the development of an industry for this high value commodity.

The income contribution of cacao to the local economy is exemplified in table 2 for the province of Aklan for 2017. The total income value contributed by cacao plantations across six municipalities is about at 4.1 million Philippine Pesos (PHP). The average exchange rate in 2017 was 59.3 PHP for 1 US Dollar. The Municipality of Libacao, with more than 12,500 trees planted and only 6,500 fruit bearing as of 2017, has the highest income value contribution in Aklan with PHP 1.4 million.

The average income to revenue ratio of a farmer planting cacao during the full term of the tree's productive maturity (5^{th} year) is 79.26% when intercropped and 71.19% when mono-cropped (Peace and Equity Foundation, 2016). Since most of the farmers in Aklan are intercropping, the estimated net income is an average maximum of 3.3 million pesos in one cropping season. This estimate of the average net income can be used to project the total property valuation of all properties planted with cacao.

Table 2, Total Economic Valuation of Cacao in Aklan.

Municipality	Productive trees as of 2017	Average production per tree (kg)*	Total value (PHP 88.80 per kg)**
Altavas	5,230	13,075	1,161,060.00
Balete	1,700	4,250	377,400.00
Ibajay	250	625	55,500.00
Libacao	6,500	16,250	1,443,000.00
Madalag	-	-	-
Malinao	4,815	12,038	1,068,930.00
Total	18,495	46,238	4,105,890.00

* A mature tree can produce an average of 2.5 kg of dried cocoa beans (Peace and Equity Foundation, 2016)
** Average farm gate price in 2011 (Peace and Equity Foundation, 2016)

With the spatial and financial information provided by this survey, decision-makers can address specific issues about the development of the cacao industry.

4. CONCLUSIONS

ODK makes data collection, analysis and mapping easy, fast and efficient. Android devices are widely available at reasonable prices and are able to perform surveys with decimetre-level accuracy. In addition, ODK solves the bottleneck of many surveys, the time-consuming and error prone manual data entry, through automatic data upload to a server database and addresses the important issue of data security if used with a dedicated server and data encryption. It is therefore suitable for fast and accurate field truthing of remote sensing projects and accuracy assessments of remotely sensed and analysed maps.

Open Data Kit is an open source software which is constantly further developed. ODK 2.0, for example, provides users with methods to build information systems on mobile devices, such as database design with tables and customised question widgets, which makes it suitable for set-up in rural locations where computers are not available. "Our own experiences combined with extensive feedback from organizations using the toolkit have led to a redesign of ODK that aims to better meet the needs of a wider range of organizations." (Brunette et al., 2013). ODK will therefore continue to meet the needs of users in a fast-changing digital world.

REFERENCES

Brunette, W., Sundt, M., Dell, N., Chaudhri, R., Breit, N., and Borriello, G., 2013. Open Data Kit 2.0: Expanding and Refining Information Services for Developing Regions. Proceedings of the 14th Workshop on Mobile Computing Systems and Application.

Chen, B., Gao, C., Liu, Y., and Sun, P., 2019. Real-time Precise Point Positioning with a Xiaomi MI 8 Android Smartphone. Sensors, 19(12): 2835.

Department of Agriculture - Bureau of Plant Industry, 2017. Philippine Cacao Industry Roadmap 2017-2022. Retrieved September 06, 2019, from http://bpi.da.gov.ph/bpi/images/PDF_file/Cacao%20Industry%20Roadmap%20-%20Signed%20%20%20March%2010,%202017.pdf

El-naggar, A.M., 2011. Enhancing the accuracy of GPS point positioning by converting the single frequency data to dual frequency data. Alexandria Engineering Journal, 50(3): 237-243.

Hartung, C., Lerer, A., Anokwa, Y., Tseng, C., Brunette, W., and Borriello, G., 2010. Open data kit: tools to build information services for developing regions. Proceedings of the 4th ACM/IEEE International Conference on Information and Communication Technologies and Development.

Pakes, U., and Subong, R., 2017. Using Andoid Devices for Smart Data Collection in the Philippines: Three case Studies from the Philippines. Philippine Quarterly of Culture & Society, 45(3&4): 107-122.

Pakes, U., Subong, R., Gumban M.A.T., and Schmitt, K., 2018. Mapping Cacao Farms and Plantations Using Android Devices. Quezon City, Philippines: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH. 70 pp. Retrieved September 06, 2019, from http://faspselib.denr.gov.ph/node/1360

Peace and Equity Foundation, 2016. A Primer on PEF's Priority Commodities: Industry Study on Cacao. Retrieved September 06, 2019, from http://pef.ph/wp-content/uploads/2016/03/Industry-Study_Cacao.pdf

Realini, E., Caldera, S., Pertusini, L., and Sampietro, D., 2017. Precise GNSS Positioning Using Smart Devices. Sensors, 17(10): 2434.

Tomkiewicz, S.M, Fuller, M.R., Kie, J.G., and Bates, K.K., 2010. Global positioning system and associated technologies in animal behaviour and ecological research. Phil. Trans. R. Soc. B, 365: 2163-2176.

van Diggelen, F., and Enge, P. 2015. The World's first GPS MOOC and Worldwide Laboratory using Smartphones. Proceedings of the 28th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2015), Tampa, Florida, pp. 361-369.

Wu, Q., Sun, M., Zhou, C., and Zhang, P., 2019. Precise Point Positioning Using Dual-Frequency GNSS Observations on Smartphone. Sensors 19(9): 2189.