A Citizen Science Approach to Classifying Urban Informality and Other Urban Land Use Types using Satellite Imagery

Tawanda Chingozha [a*], Dieter Von Fintel [a]

[a] *Economics Department, Stellenbosch University, PRIVATE BAG X1, 7602 MATIELAND, SOUTH AFRICA;*

Abstract

Developing countries are characterised by rapid and often unplanned urban expansion. While Hegazy and Kaloop (2015) suggest that urban sprawl into the hinterland has long been associated with economic vibrancy. This phenomenon is also a consequence of very high rates of rural-urban migration, and it might result in environmental problems as service delivery capacity fails to cope with the unplanned nature of the expansion. The unplanned nature and unregistered status of residential and commercial expansion respectively, limits economic potential because the lack of land titles (in unplanned settlements) and registration (for informal businesses) is characterised by lack of trust between informal settlement communities and informal business owners; and government. Urban land cover mapping can reveal information on the size and growth patterns of informal businesses and settlements and is an important beginning towards understanding their role in economic vitality and growth. The exercise may help establish cadastral databases that may then be used in formalising these areas and providing essential services. This study explores citizen science as an approach for classifying urban informality and other land use types, given the benefits of such information in urban land use planning, cadastral database development and the need to understand the role played by urban unplanned residential and commercial expansion in economic vitality.

Key Words: citizen science, land cover mapping, urban expansion, economic vitality, urban informality

1.0 Introduction

Land cover as “the biophysical cover of the Earth’s terrestrial surface, identifying vegetation, inland water, bare soil or human infrastructure” (Chen, Li, Wu, & Chen, 2017; Gómez, White, & Wulder, 2016). Di Gregorio (2005), comments that land cover and its alterations is an expression of human activity. It is a critical environmental variable, key in modelling economic land use and environmental monitoring and assessments (C. Fonte, L. Bastin, L. See, G. Foody, & F. Lupia, 2015; Fonte et al., 2017; Foody et al., 2013; Fritz et al., 2009; Fritz et al., 2017). The mapping and analysis of land cover is an important step in the natural resource management and studies that seek to understand the distribution of habitats (Gómez et al., 2016). Information on land use types is hence a critical strategy in tracing changes in informal housing and informal businesses in developing countries. Developing countries bring to the fore the importance of land cover mapping for urban management and sprawl management (Hegazy & Kaloop, 2015; Malarvizhi, Kumar, & Porchelvan, 2016; See et al., 2015; Vaz & Jokar Arsanjani, 2015) because they are characterised by unorganized and unplanned urban expansion. Vaz and Jokar Arsanjani (2015), acknowledge the important role played by urban regions and major metropolis in future economic stability. However, this relies on the ability for cost-effective and accurate mapping and development of cadastral data, especially for developing countries. Urban cadastral data obtaining from land cover maps can assist in service provision and land title provision in informal settlements for example. Against this and the background of and “African statistical tragedy” [see Devarajan (2013)], this study investigates the citizen science approach for the classification of the informal sector and other urban land use types.

Advances in web technology have allowed volunteers to be at the forefront in producing different kinds of geographic data (Arsanjani & Vaz, 2015; Chapman, Bell, & Bell, 2017). This is known as citizen science, crowdsourcing, human computation or Volunteered Geographic Information (VGI) (Albuquerque, Herfort, & Eckle, 2016; Arsanjani & Vaz, 2015; C. C. Fonte, L. Bastin, L. See, G. Foody, & F. Lupia, 2015). User contributions in VGI projects tend to involve tracing out specific land use types on a georeferenced image, while in other cases this may involve the collection of GPS points by the users (Arsanjani & Vaz, 2015; C. C. Fonte et al., 2015; Neis & Zielstra, 2014). This study relied on the former approach (classification), and it involved users tracing out different land use types on georeferenced Google Earth (GE) imagery uploaded on the Zooniverse platform. Classification is the “process of assigning predefined attributes (values/categories) to existing geographical information” and is the first genre of analytical task that participants in a crowdsourcing project can undertake and it is associated with the deployment of interpretive skills and past/background information to classify the piece of geographic information at hand (Albuquerque et al., 2016). Hence, background information and the predefinition of values is an important aspect in citizen science project. The study made use of the Zooniverse[1] platform.

To test whether the development of an informality dataset (and other land use types) using satellite imagery and citizen science is the overarching goal of this paper. A secondary objective is to quantify the magnitude and extent of the Zimbabwe's 2005 clean-up operation named Operation Restore Order (ORO) or Operation "*Murambatsvina*[2]", the outcome of which has the potential to be linked with other socio-economic data. Most importantly though, ORO allows us to test whether citizen science is able to detect changes in informal land use through including pre (2004) and post (2006) exercise images for areas affected and those that were not. Hence, Section 2 presents some background of the informal sector and other urban land use types. Section 3 discusses some issues related to crowdsourcing, as well as the experimental design. Section 4 presents the results from the study while conclusions and policy recommendations are presented in Section 5.

2.0 Urban Informality

2.1 Some theoretical Overview

Falco, Kerr, Rankin, Sandefur, and Teal (2011) and Rothenberg et al. (2016) define the informal sector as all economic activities[3] that neither pay taxes nor are registered with the government. ILO (1972), Rei and Bhattacharya (2008) and Batini, Levine, Kim, and Lotti (2010) describe informal services as activities that are unrecorded, unrecognized, unprotected and unregulated by the authorities. According to Benjamin, Mbaye, and Diop (2012) the informal sector is the group of small and organized producers who operate on the boundaries of the formal economy, while Becker (2004) views the informal sector as the unregulated and non-formal element of the market economy that produces goods and services for other forms of remuneration.

[1] www.zooniverse.org
[2] Murambatsvina means "away with filth" in vernacular Shona.
[3] This is our preferred definition because the informal can also include other non-business entities such as informal settlements

Literature offers diverse views on the factors behind informality. Bhattacharya (1995), stresses that dual economy models of development disproportionately enlarge the role of agriculture, yet ancient societies were engaged in some kind of crafts-making – which means informality has always been a permanent component of the industrialisation process. Despres (1985), asserts that informality is a consequence of the absence of an objective relationship between high profits and low wages in the formal sector. Hart (1973), explains that the informal sector is influenced by the involuntary need to reduce the persistent imbalance between wage employment income and expenditure, while Rei and Bhattacharya (2008) view the informal sector as a response by economic agents to over-regulation by the government. Batini et al. (2010), views the informal sector as a natural occurrence in countries with a huge tax burden and weak enforcement mechanisms such that there is a negative relationship between economic performance and informality. The most common cause, however, is that the formal sector cannot absorb all labour force participants (Becker, 2004; Benjamin et al., 2012; Despres, 1985; Naik, 2009).

There are divergent views on the relationship between the economic performance of the formal sector and the informal sector; some hold pro-cyclical and others anti-cyclical views (Gerxhani, 2004). The anti-cyclical effect occurs when labour shifts into informal employment as the formal economy contracts (a substitution between sectors). The pro-cyclical effect dominates when the direct and indirect demand of the products and services of the informal sector expands as the formal economy registers growth (complementarity between sectors). The informal sector is, however, in the majority of cases a product of economic stagnation of the formal sector (Batini et al., 2010; Gerxhani, 2004; Luebker, 2008). Shapiro (2015), provides a different view, explaining that the relationship between informality and economic performance hinges on the institutions in the economy.

Verick (2006), assert that the informal sector accounts for 78% of employment in sub-Saharan Africa excluding South Africa, indicating potentially high levels of competition parts of the formal sector. In contrast, Becker (2004) gives a lower figure, claiming that informal employment actually accounts for 70% of jobs in sub-Saharan Africa. The figures are slightly lower for Latin America and Asia at 60% and 59% respectively (Becker, 2004). Benjamin et al. (2012), estimate that between 80 – 90% of employment in Africa as a whole is in the informal sector, and that up to 60% of national income in West Africa is derived from informal activities. Clearly, the informal sector has significance in developing countries and that is why of late there is renewed interest on the subject beyond economics (Naik, 2009).

There is little known about the urban informality (size, growth patterns and GDP contribution) and there has been a general neglect of the informal sector from the policy, national income accounting and research perspectives despite its importance (Benjamin et al., 2012; Gerxhani, 2004; Onwe, 2013). Therefore, more understanding of the informal sector adds momentum to poverty alleviation through support.

2.2 Measurement of Urban Informality

Urban informality (both business and housing) forms a significant part of the economy in developing countries. The sector is integral to the livelihood and sustenance of particularly the poor, and it normally acts as the first point of employment for people migrating from rural to urban areas (Bhattacharya, 1995). The informal sector sometimes becomes an innovation hub, and policy towards informality should not be from the attitude that it contributes to urban unemployment but to recognise its potential (Bhattacharya, 1995). Despite all these positives, little is understood about the informal sector. By definition, the sector is unregulated: therefore obvious sources of data from banking, tax returns and other activities are generally absent. Often governments do not even know the size of the informal sector and its monetary contribution; the result is negative state action in the form of clean-up operations against the informal economy. This study examines the feasibility of using satellite imagery and citizen science in the classification of urban informality and other land use types.

2.3 Urban Informality in Zimbabwe

Informality is a natural phenomenon for Zimbabwe. Historically, its economy has not been able provide enough employment for the population (Ncube, 2000). Even before independence, this problem was significant in the 1970s as British sanctions and instability due to the liberation war hurt the Rhodesian economy. Growth after independence was erratic, although the public sector managed to play a critical role in absorbing large numbers of people (Ncube, 2000). However, The Economic Structural Adjustment Program (ESAP) called for the laying off redundant workers in the civil service. As a result, the absorptive capacity of the public service was negatively affected and people were left with no choice but to join the informal sector.

According to Tibaijuka (2005), at independence the informal sector accounted for less than 10% of employment, but grew to 40% by 2005 due to the gradual economic decay that Zimbabwe was experiencing. Thus, Operation Restore Order of 2005 represented a direct negative shock that affected 40% of the employed labour force in 2005. Before ESAP the living standards of Zimbabwe's urbanites were probably the best in sub-Saharan Africa (Potts, 2006), thus the post-ESAP period is important in explaining informality. Zimbabwe's informal sector enterprises are small, employing less than five people (Rothenberg et al., 2016). The study also considers informal housing, mainly cooperative housing because in most cases, plot owners settle on the land before the required amenities such as sewerage infrastructure, paved roads and piped water have been put in place due to pressure on the housing market.

2.4 The 2005 Clean-up Operation in Zimbabwe

Clean-up campaigns have proven to be an integral part of social policies in a number of developing countries: for example, the 2005 "Operation Restore Order" in Zimbabwe (Potts, 2006; Tibaijuka, 2005), the 2006 "Operation Dongosolo" in Malawi (Riley, 2014) and the demolition of Muoroto and Mwariro illegal settlements in Kenya in the 1990s (Arimah & Branch, 2011). The absence of informal sector recognition may be an influencing factor behind clean-up campaigns and other negative state action for example. The national and by-laws that are used by African governments are often considered to be out-dated and to represent fragments of a colonial inheritance that the political establishment retain to exploit the masses (Arimah & Branch, 2011; De Soto, 2000; Potts, 2006; Riley, 2014; Tibaijuka, 2005). As discussed by Hegazy and Kaloop (2015) and Malarvizhi et al. (2016), unplanned urban expansion is responsible for several environmental issues in developing countries. At the same time, Hegazy and Kaloop (2015) maintain that urban sprawl into the rural hinterland (both for residential and commercial purposes) has long been regarded as the hallmark of economic vitality of the region.

Operation Restore Order was unprecedented, and according to Potts (2006), there is no other recorded urban eviction program to sweep across the metropolitan areas of entire cities on the African continent, including Apartheid South Africa. It was launched on the 25th of May 2005 in Harare, the capital of Zimbabwe, quickly spreading into other cities and towns, and by the end about 700,000 people had been rendered homeless or without a source of livelihood. Government sought to quell all forms of illegality in the form of unregistered business and residential structures (Tibaijuka, 2005). According to Potts (2006), government put its own estimate of the total number of people of affected at 570,000, with 92,460 dwellings being destroyed, 133,534 households losing their places of abode and 98,000 losing their sources of income. As a result of the clean-up campaign poverty, destitution, deprivation and vulnerability increased (Potts, 2006).

A secondary aim of the study is to buttress the survey findings of Tibaijuka (2005) and others with objective, indisputable satellite data from the sample of Harare under study. Importantly though, ORO is an real event that allows us to test whether users in crowdsourcing projects can detect changes in urban informality and other land use types through the inclusion of pre and post clean-up exercise images in the experiment.

3.0 The Citizen Science Approach

The popularity of citizen science and VGI in the mapping of urban land use types is not without merit. Albuquerque et al. (2016), asserts that crowdsourcing may be a superior approach to land use classification where features are heterogeneous and inconsistent. The informal sector in Zimbabwe is one such area of study where it takes different shapes and forms – potentially rendering the crowdsourced approach more useful than automatic detection.

In the validation of land cover maps, C. Fonte et al. (2015) explains that physical ground-truthing is laborious, may be impossible if researchers are far away from the areas of study and may suffer the drawback of the ground-based experts not agreeing on the classifications. VGI datasets and classifications can be used as reference data, resulting in efficiency and cost savings. An important channel through which cost savings is the availability of free geo-referenced JPEG format images available from Google Earth (GE) platform, given that multi-spectral band information is usually unnecessary. Malarvizhi et al. (2016), note that GE provides free VHR imagery that is suitable for visualization and the image-based classification that crowdsourcing provides and it provides good temporal resolution.

The GE platform offers critical functionality for the validation of global land cover maps (Fritz et al., 2009), hence it was an important source of imagery for the study. The GE platform hosts VHR imagery with a good resolution such as 50cm x 50cm, which allows even non-remote sensing experts to classify different land use classes in a reliable and cost effective manner (Fritz et al., 2009). Following Malarvizhi et al. (2016), the images were downloaded and georeferenced using Elishayal Smart GIS tool.

3.1 Past Examples of Citizen Science projects

There have been a number of papers that have applied crowdsourcing techniques in the measurement of land use, cultural and socio-economic changes [see Pettorelli, Gliozzo, and Haklay (2016)]. In terms of web applications, Open Street Map (OSM) is among the most popular VGI projects that the academic and research community has had a keen interest in (Fonte et al., 2017). Neis and Zielstra (2014), posit that the objective of the OSM project is to establish an open source geographic information database for use in mapping, navigation and other case uses. Another one is the Geo-Wiki project. It was established the promote global networking by volunteers with the end goal of enhancing global land cover maps through crowdsourcing (Foody et al., 2013; Fritz et al., 2017). The Zooniverse platform (the one that this study relies on) was developed by astronomers and physicists to aid the identification of new galaxies from thousands of night time images of the sky taken by telescopes. Since then, the platform has hosted a plethora of different image classification projects from different fields.

3.2 Some Drawbacks of the Citizen Science Approach

Despite the numerous advantages of citizen science, there a few drawbacks. Explaining the importance of VGI data in land use mapping, Dorn, Törnros, and Zipf (2015) and Foody et al. (2013) acknowledge that the fact that the volunteers generating the data are untrained in an important concerned. Following Fritz et al. (2017)'s idea on the need to develop a training manual, this study create video and in-application/on-demand tutorials in order to help volunteers on the projects to visualise the predefined typology and have some background information; and to foster training.

Apart from the need for training, class noise, participant inequality and time-consuming annotations are additional problematic issues in VGI/crowdsourcing. Explaining some of the inherent issues with Open Street Map (OSM) data, Johnson and Iizuka (2016) and Johnson, Iizuka, Bragais, Endo, and Magcale-Macandog (2017) highlight the problem of class noise in the data that are caused by classification errors on the part of the users. Small objects (potentially the identification of informal structures such as shacks) can exacerbate class noise. Albuquerque et al. (2016), assert that there is little likelihood for volunteers to classify small objects correctly. Such issues are important to consider in the classification of the informal sector and other land use types. Class noise is tested in this study through the calculation of classification accuracy rates.

Participant inequality is a critical issue that characterise VGI platforms (Neis & Zielstra, 2014; Nielsen, 2006; Stewart, Lubensky, & Huerta, 2010). Known as the 90-9-1 rule, this is a scenario is which 90% of users on a VGI project never contribute anything; 9% having an irregular presence on the platform and 1% accounting for almost all the user contributed data on the project (Nielsen, 2006). As our results show, this problem affected the structure of the experiment substantially as shown in Section 4. VGI projects also do take time. Relative to machine learning, the classification of imagery using citizen science is essentially based on photo interpretation and it is time-consuming (Srivastava, Lobry, Tuia, & Vargas-Muñoz, 2018; Xing, Meng, Hou, Song, & Xu, 2017), although this is an important advantage that relying on a bigger 'crowd' of volunteers is supposed to solve.

3.3 Experiment Setup and Implementation

The study employed 180 images split equally between 2004 and 2006 for the same areas (see Figure 1). The images were randomly allocated into four streams as shown in Table 1 contain Images for areas that were not affected by ORO were allocated to Stream 1 and 2, hence the images look the same for both years. Images for areas affected by ORO were allocated to Streams 3 and 4; hence, the major difference between the 2004 and 2006 images are that informal housing and informal business land use types were destroyed during the 2005 clean up exercise. The experiment was hence set-up to test whether citizen science can detect changes in urban informality and other land use types. As explained by Hegazy and Kaloop (2015), monitoring urban growth using RS data involves detecting changes between two periods that is uncharacteristic of normal variation. ORO allows us to do exactly that.

[INSERT FIGURE1 HERE]

Table 1 Experiment Design

	No. of images			Classifications			
Year	2004	2006	Totals	Per image	2004	2006	Totals
Stream 1 - No change	25	25	50	5	125	125	250
Stream 2 - No change	20	20	40	10	200	200	400
Stream 3 - Change	25	25	50	5	125	125	250
Stream 4 - Change	20	20	40	10	200	200	400
Totals	90	90	180				1300

Foody et al. (2013), C. Fonte et al. (2015) and Albuquerque et al. (2016) indicate that one way of improving the quality of data generated via citizen science and to reduce class noise is to rely on consensus and agreement amongst many volunteers for the same piece of work or area. Albuquerque et al. (2016), found that tasks with the highest level of agreement were 41 times more likely to be classified correctly in a crowdsourced project relative to those with low levels of agreement. The design of the pilot study for Stream 2 and Stream 4 (such that an image is classified by 10 users, instead of 5) centred on 'Linus' Law. Linus Law is the supposition that there is a positive correlation between the number of contributors and quality, thereby providing an intrinsic measure of quality assurance (Haklay, Basiouka, Antoniou, & Ather, 2010). This was a key factor under investigation in the experiment design.

We ran a student recruitment campaign on campus and forty-one volunteers signed up for the project. Albuquerque et al. (2016), recognise the need for citizen science projects to collect data on participant skills prior to the image classification work in order to inform the allocation of users to undertake specific tasks. As Foody et al. (2013) put across; volunteers may vary from eager but naïve and untrained individuals to those who are highly experienced and skilled. During the recruitment phase, students were asked to fill in a questionnaire to collect some basic demographic information; not for the basis of selecting individuals into different groups by perceived skills [for example, following Meier (2013)] but in order to check whether the different groups under the experiment were statistically balanced.

Table 2 Allocation of students to streams

Stream	Number of students	Images/student
1	8	31.25
2	12	33.33
3	9	27.78
4	12	33.33

Table 2 shows the number of students out of the 41 allocated to each stream and the number of images each student needed to classify to ensure that the experiment worked and to prevent participant inequality; while Table 3 shows the results of the balance tests. Fischer's Exact Test checks whether different treatments results in different outcomes in an experiment. In this case, the different treatments are the different individuals making up the four streams. Under Fischer's Exact Test, the null hypothesis (treatment affects outcome) is not rejected if p is small. Table 3 shows that the allocation of participants into stream was balanced.

Table 3 Fischer Tests for Statistical Balance

Variable	Fischer Test p-value
Frequently busy from informal business?	0.31
Informal business customer?	1.00
Owns informal business?	0.75
Frequently visits informal settlement?	0.38
Visited informal settlement?	0.12
Lived informal settlement?	0.66
Household Income	0.65
Education	0.31
Parent Education	0.10
Sex	0.77
Age	0.26
Province	0.96
Nationality	0.37
Race	0.97

There is no clear knowledge about the specific tasks most suitable for classification using citizen science (Albuquerque et al., 2017). In this study, participants were presented with five tasks[4], as shown in Table 4. The tasks involved the identification of four different land use types namely a) informal businesses, b) informal housing/cooperative housing, c) low income formal housing, d) high income formal housing and e) formal businesses.

Table 4 Classification Tasks

Task Number	Task	Instruction on platform
a	informal businesses	Draw around the area(s) that appear to be informal businesses or business hubs.
b	informal housing/cooperative housing	Draw around the area(s) that appear to be informal housing or cooperative settlements.
c	low income formal housing	Draw around the area(s) that appear to be lower income formal housing.
d	high income formal housing	Draw around the area(s) that appear to be higher income formal housing.
e	Formal businesses	Draw around the area(s) that appear to be formal businesses or industry

The participants completed the tasks by tracing polygon(s) around specific land use types in accordance with the task instructions on the platform. The platform carried clear instructions

[4] https://www.zooniverse.org/projects/tchingozha/identifying-and-measuring-the-urban-informal-economy-and-other-land-use-types

for participants to "do nothing" if they could not identify particular land use classes on an image. Participants traced out the different land use types using a Bezier tool (see Figure 2). Tutorials to distinguish the different land use types, and how to navigate the interface were inbuilt within the platform while videos were made available on YouTube as stated earlier in the paper.

Figure 2 Classification using the Bezier tool

3.4 Processing Classification Results

The classification output from the project was exported from the Zooniverse platform as Excel files for each stream. The files contained several columns of information such as participant user names, image names, task numbers but mostly importantly the X and Y map coordinates for redrawing the user classification and reference[5] polygons. The Excel files were prepared in STATA for eventual plotting and further 'cleaning' in Quantum GIS (QGIS) software. Some of the polygons drawn by users violated some geometry properties of ESRI shape files (in some instances polygon lines crossed as shown in Figure 3). Such instances were corrected by

[5] The authors did an expert (reference) classification of the all the images in the project, based on their familiarity and knowledge of the areas sampled out of Harare for the study – together with their knowledge of the classification typology

manually retracing the polygons, making sure that lines do not cross through making multi-parts out of the erroneous polygon (from a geometry properties perspective).

Figure 3 Creating multi-parts from single polygon to address crossing lines

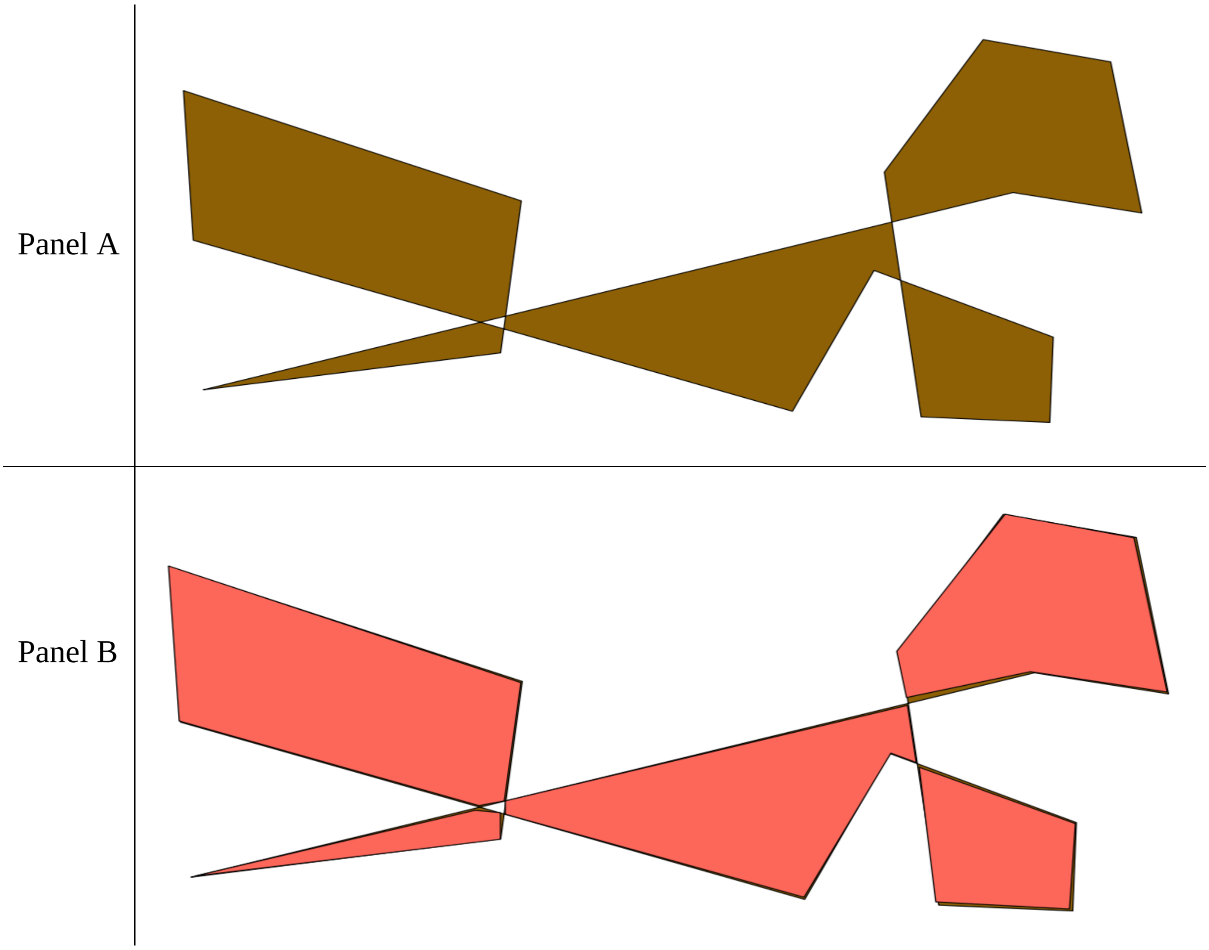

Figure 3's panel A shows a polygon with crossing lines, and panel B shows correction by splitting the polygon into multi-parts. Aside from the manual correction shown in Figure 3, the rest of geo-processing: i) plotting the XY coordinates of each classification, ii) converting points to polygons, iii) intersecting user and reference classifications and iv) calculating the areas of overlap was done programmatically using PYQGIS. The data was exported back to STATA for the final analysis.

4.0 Results and Discussion

Quality and accuracy is a key issue in VGI (Fritz et al., 2009). C. Fonte et al. (2015), note that the accuracy of a land cover map is usually derived from the extent of its agreement to some gold standard of ground truth. We did an expert classification of the area under study given our in-depth knowledge and familiarity with it as already mentioned. Classification accuracy was calculated using the formula:

$$Classification\ Accuracy = \frac{Overlap\ Area}{Reference\ \ Area}$$

The classification accuracy rates are shown in Table 5. The low accuracy rates shown in Table 5 potentially signal the need for better participant training methods. Fritz et al. (2017), note that the Geo-Wiki platform may be used to train participants in crowdsourced projects. However, a potential setback may be that the platform might not have more coverage of developing countries where informality is a huge phenomenon.

Table 5 Classification Accuracy Rates

Stream	Classifications	Mean	Std. Dev.	Min	Max
1	129	0.62	0.2683	0.03865	1
2	95	0.58	0.30335	0.00009	1
3	67	0.65	0.28746	0.00039	1
4	62	0.62	0.32437	0.00312	1

The study also noted that some significant variation in accuracy by class, which means that participants found some classes difficult to classify that others. This finding concurs with Foody et al. (2013) who also find class-specific difference in quality of information obtained via VGI.

The results in Table 6 show that the 2005 clean-up operation in Zimbabwe reduced the area covered by informal businesses and backyard structures within formal housing. There is a positive growth in formal housing which may be a substitution effect after the demolition of informal residential areas. Formal businesses show stagnation, which means that the destruction of informal business structures did not translate to increased formal commercial activities.

Table 6 Multinomial Logit Estimates

Dep. Var: Area	Beta/(S.E)
Informal business	-1.55
	(0.2489)***
Backyard structures	-5.12
	(1.0060)***
Formal housing	0.407
	(0.1252)***
Formal business	-0.132
	(0.1449)***
N	1747
R^2	0.0565

NOTES: The multinomial regression estimates area as a function of year (2004 as base year) and land use type.

5.0 Conclusion and Recommendations

Fairbairn and Al-Bakri (2013), maintain that it is a much better scenario to have "no mapping at all" rather than have rely on inaccurate and inconsistent VGI generated data. Considering the relatively low classification accuracy rates generated in this study, it is clear that identification of informality and other land use types through citizen science has a few shortcomings. From the results, there are a few strategies that could potentially be implemented to better the quality of the results in the future. These findings and conclusions are important contribution on their own, but from the reference classification, there are few comments which are possible to make. The multinomial logit estimates confirm the scale of the destruction brought about by ORO on the Zimbabwe's urban landscape, which reinforces findings by Tibaijuka (2005) as well what the Afrobarometer 2005 survey showed.

References

Albuquerque, J., Herfort, B., & Eckle, M. (2016). The tasks of the crowd: A typology of tasks in geographic information crowdsourcing and a case study in humanitarian mapping. *Remote Sensing, 8*(10), 859.

Arimah, B. C., & Branch, C. M. (2011). *Slums as expressions of social exclusion: Explaining the prevalence of slums in African countries.* Paper presented at the OECD International Conference on Social Cohesion and Development, Paris.

Arsanjani, J. J., & Vaz, E. (2015). An assessment of a collaborative mapping approach for exploring land use patterns for several European metropolises. *International Journal of Applied Earth Observation and Geoinformation, 35*, 329-337.

Batini, N., Levine, P., Kim, Y.-B., & Lotti, E. (2010). Informal labour and credit markets: a survey. *IMF Working Papers*, 1-41.

Becker, K. F. (2004). *The informal economy: fact finding study*: Sida.

Benjamin, N., Mbaye, A. A., & Diop, I. T. (2012). *The informal sector in Francophone Africa: firm size, productivity, and institutions*: World Bank Publications.

Bhattacharya, P. C. (1995). The economics of development: A review article. *Journal of Economic Studies, 22*(2), 59-74.

Chapman, L., Bell, C., & Bell, S. (2017). Can the crowdsourcing data paradigm take atmospheric science to a new level? A case study of the urban heat island of London quantified using Netatmo weather stations. *International Journal of Climatology, 37*(9), 3597-3605.

Chen, J., Li, S., Wu, H., & Chen, X. (2017). Towards a collaborative global land cover information service. *International journal of digital earth, 10*(4), 356-370.

De Soto, H. (2000). *The mystery of capital: Why capitalism triumphs in the West and fails everywhere else*: Basic books.

Despres, L. A. (1985). *Industrialization, Migration and the Informal Sector in Manaus*: University of Notre Dame, Helen Kellogg Institute for International Studies.

Devarajan, S. (2013). Africa's statistical tragedy. *Review of Income and Wealth, 59*(S1).

Di Gregorio, A. (2005). *Land cover classification system: classification concepts and user manual: LCCS* (Vol. 2): Food & Agriculture Org.

Dorn, H., Törnros, T., & Zipf, A. (2015). Quality evaluation of VGI using authoritative data—A comparison with land use data in Southern Germany. *ISPRS International Journal of Geo-Information, 4*(3), 1657-1671.

Fairbairn, D., & Al-Bakri, M. (2013). Using geometric properties to evaluate possible integration of authoritative and volunteered geographic information. *ISPRS International Journal of Geo-Information, 2*(2), 349-370.

Falco, P., Kerr, A., Rankin, N., Sandefur, J., & Teal, F. (2011). The returns to formality and informality in urban Africa. *Labour Economics, 18,* S23-S31. doi: 10.1016/j.labeco.2011.09.002

Fonte, C., Bastin, L., See, L., Foody, G., & Lupia, F. (2015). Usability of VGI for validation of land cover maps. *International Journal of Geographical Information Science, 29*(7), 1269-1291.

Fonte, C., Minghini, M., Patriarca, J., Antoniou, V., See, L., & Skopeliti, A. (2017). Generating up-to-date and detailed land use and land cover maps using OpenStreetMap and GlobeLand30. *ISPRS International Journal of Geo-Information, 6*(4), 125.

Fonte, C. C., Bastin, L., See, L., Foody, G., & Lupia, F. (2015). Usability of VGI for validation of land cover maps. *International Journal of Geographical Information Science, 29*(7), 1269-1291.

Foody, G. M., See, L., Fritz, S., Van der Velde, M., Perger, C., Schill, C., & Boyd, D. S. (2013). Assessing the accuracy of volunteered geographic information arising from multiple contributors to an internet based collaborative project. *Transactions in GIS, 17*(6), 847-860.

Fritz, S., McCallum, I., Schill, C., Perger, C., Grillmayer, R., Achard, F., . . . Obersteiner, M. (2009). Geo-Wiki. Org: The use of crowdsourcing to improve global land cover. *Remote Sensing, 1*(3), 345-354.

Fritz, S., See, L., Perger, C., McCallum, I., Schill, C., Schepaschenko, D., . . . Laso-Bayas, J.-C. (2017). A global dataset of crowdsourced land cover and land use reference data. *Scientific Data, 4*, 170075.

Gerxhani, K. (2004). The informal sector in developed and less developed countries: a literature survey. *Public choice, 120*(3-4), 267-300.

Gómez, C., White, J. C., & Wulder, M. A. (2016). Optical remotely sensed time series data for land cover classification: A review. *ISPRS Journal of Photogrammetry and Remote Sensing, 116*, 55-72.

Haklay, M., Basiouka, S., Antoniou, V., & Ather, A. (2010). How many volunteers does it take to map an area well? The validity of Linus' law to volunteered geographic information. *The Cartographic Journal, 47*(4), 315-322.

Hart, K. (1973). Informal income opportunities and urban employment in Ghana. *The journal of modern African studies, 11*(01), 61-89.

Hegazy, I. R., & Kaloop, M. R. (2015). Monitoring urban growth and land use change detection with GIS and remote sensing techniques in Daqahlia governorate Egypt. *International Journal of Sustainable Built Environment, 4*(1), 117-124.

Johnson, B. A., & Iizuka, K. (2016). Integrating OpenStreetMap crowdsourced data and Landsat time-series imagery for rapid land use/land cover (LULC) mapping: Case study of the Laguna de Bay area of the Philippines. *Applied Geography, 67*, 140-149.

Johnson, B. A., Iizuka, K., Bragais, M. A., Endo, I., & Magcale-Macandog, D. B. (2017). Employing crowdsourced geographic data and multi-temporal/multi-sensor satellite imagery to monitor land cover change: A case study in an urbanizing region of the Philippines. *Computers, Environment and Urban Systems, 64*, 184-193.

Luebker, M. (2008). Employment, unemployment and informality in Zimbabwe: Concepts and data for coherent policy-making.

Malarvizhi, K., Kumar, S. V., & Porchelvan, P. (2016). Use of high resolution google earth satellite imagery in landuse map preparation for urban related applications. *Procedia technology, 24*, 1835-1842.

Meier, P. (2013). Human computation for disaster response *Handbook of human computation* (pp. 95-104): Springer.

Naik, A. K. (2009). *Informal sector and informal workers in India.* Paper presented at the seminar measuring the informal economy in developing countries, Kathmandu, Nepal.

Ncube, M. (2000). Employment, unemployment and the evolution of labour policy in Zimbabwe. *Zambezia, 27*(2), 161-194.

Neis, P., & Zielstra, D. (2014). Recent developments and future trends in volunteered geographic information research: The case of OpenStreetMap. *Future internet, 6*(1), 76-106.

Nielsen, J. (2006). Participation inequality: Encouraging more users to contribute. *http://www. useit. com/alertbox/participation_inequality. html*.

Onwe, O. J. (2013). Role of the Informal Sector in Development of the Nigerian Economy: Output and Employment Approach. *J Econ Dev Stud, 1*, 60-74.

Pettorelli, N., Gliozzo, G., & Haklay, M. (2016). Using crowdsourced imagery to detect cultural ecosystem services: a case study in South Wales, UK. *Ecology and Society, 21*(3).

Potts, D. (2006). 'Restoring Order'? Operation Murambatsvina and the Urban Crisis in Zimbabwe. *Journal of Southern African Studies, 32*(2), 273-291. doi: 10.1080/03057070600656200

Rei, D., & Bhattacharya, M. (2008). *The impact of institutions and policy on informal economy in developing countries: an econometric exploration*: Citeseer.

Riley, L. (2014). Operation Dongosolo and the geographies of urban poverty in Malawi. *Journal of Southern African Studies, 40*(3), 443-458.

Rothenberg, A. D., Gaduh, A., Burger, N. E., Chazali, C., Tjandraningsih, I., Radikun, R., . . . Weilant, S. (2016). Rethinking Indonesia's informal sector. *World Development, 80,* 96-113.

See, L., Fritz, S., Perger, C., Schill, C., McCallum, I., Schepaschenko, D., . . . Kraxner, F. (2015). Harnessing the power of volunteers, the internet and Google Earth to collect and validate global spatial information using Geo-Wiki. *Technological Forecasting and Social Change, 98,* 324-335.

Shapiro, A. F. (2015). Institutions, Informal Labor Markets, and Business Cycle Volatility. *Economía, 16*(1), 77-112.

Srivastava, S., Lobry, S., Tuia, D., & Vargas-Muñoz, J. (2018). *Land-use characterisation using Google Street View pictures and OpenStreetMap.* Paper presented at the Proceedings of the Association of Geographic Information Laboratories in Europe Conference (AGILE).

Stewart, O., Lubensky, D., & Huerta, J. M. (2010). *Crowdsourcing participation inequality: a SCOUT model for the enterprise domain.* Paper presented at the Proceedings of the ACM SIGKDD Workshop on Human Computation.

Tibaijuka, K. (2005). UN human special envoy on human settlements issues in Zimbabwe: Report of the fact-finding mission to Zimbabwe to assess the impact of operation Murambatsvina. *Vienna: UN Publications*.

Vaz, E., & Jokar Arsanjani, J. (2015). Crowdsourced mapping of land use in urban dense environments: An assessment of Toronto. *The Canadian Geographer/Le Géographe canadien, 59*(2), 246-255.

Verick, S. (2006). The impact of globalization on the informal sector in Africa. *Economic and Social Policy Division, United Nations Economic Cat Commission for Africa*.

Xing, H., Meng, Y., Hou, D., Song, J., & Xu, H. (2017). Employing crowdsourced geographic information to classify land cover with spatial clustering and topic model. *Remote Sensing, 9*(6), 602.