Study on the accuracy of school location information in South Africa

Lauren Hankel [1 2], Melissa Burgess [2], Kobus Roux [2], Anita van Deventer [2], Merryl Ford [2], Ronel Smith [2], Sives Govender [2]

[1] Meraka Institute, Council for Scientific and Industrial Research, Pretoria, South Africa, lhankel@csir.co.za

[2] Meraka Institute, Council for Scientific and Industrial Research, Pretoria, South Africa

DOI: http://dx.doi.org/10.4314/sajg.v6i2.*1*

Abstract

Accurate location information is required for proper planning and informed decision making in a variety of sectors. In the basic education sector, accurate school location information is typically required for road, electricity, internet connectivity and water infrastructure planning as well as planning for the delivery of textbooks and public transport (i.e. busses, taxis). The National Education Collaboration Trust (NECT) commissioned the Council for Scientific and Industrial Research (CSIR) to conduct a study of existing school location information in five NECT education districts (Bohlabela, Bojanala, Mount Frere, Uthungulu and Waterberg). School location information in two existing databases, i.e. Education Management Information System (EMIS) and National Education Information Management System (NEIMS) were assessed. Due to the nature of school buildings (i.e. varying form of schools) it is challenging to automate the identification of schools from satellite imagery by using machine learning/image processing techniques. Manual Geographic Information System (GIS) techniques were applied to conduct the study. High resolution satellite imagery and Google StreetView were utilised to ascertain the locations of schools. This study indicated that there are discrepancies between the EMIS and NEIMS databases and that there is a significant amount of school location information that might not be useful for proper planning and informed decision making in certain sectors due to the degree of positional inaccuracy of the data. If the positional accuracy of the incorrect school location information improves, it will have a positive impact on the overall outcomes of planning and decision making.

1. Introduction

Accurate location information is required for proper planning and informed decision making in a variety of sectors, therefore, accurate school location information is required for planning and decision making in education. The Department of Basic Education (DBE) is the custodian of information on all schools in South Africa. A data custodian is defined as “(a) an organ of state" (Spatial Data Infrastructure Act, 2003), or "(b) an independent contractor or person engaged in the exercise of a public power or performance of a public function, which captures, maintains, manages, integrates, distributes or uses spatial information" (Spatial Data Infrastructure Act, 2003). Education Management Information Systems (EMIS) is a function in DBE that is responsible to develop and maintain an integrated education information system for education management (Department of Basic Education, 2016). DBE and provincial EMIS units have a mandate to provide education information to the education system, and to support monitoring, planning and decision-making processes (Department of Basic Education, 2014). DBE and the Provincial Education Departments conduct surveys on a yearly basis. School location information are typical outputs of these surveys.

2. Background

In order to provide infrastructure services to schools, accurate location information is critical. One of the largest projects currently being undertaken is that of providing broadband connectivity to schools. The South Africa (SA)-Connect initiative of the Department of Telecommunications and Postal Services (DTPS) is the vehicle for implementing South Africa’s Broadband Policy, and is starting to rollout connectivity to schools throughout the country (Department of Telecommunications and Postal Services, 2015). Phase one of SA-Connect aims to provide broadband access of 10 megabytes per second (Mbps) to 4444 schools in eight District Municipalities by the end of 2017. Targets increase incrementally towards one gigabyte per second (Gbps) to all schools by 2030 (Department of Telecommunications and Postal Services, 2015).

The National Education Collaboration Trust (NECT) is an organisation that is dedicated to strengthening collaboration between civil society and government in order to achieve the basic education goals that are listed in the 2030 National Development Plan (NDP) (Republic of South Africa, 2012) of South Africa (National Education Collaboration Trust, 2017). In 2016 NECT commissioned the Council for Scientific and Industrial Research (CSIR) to conduct a study of existing school location information for connectivity planning in five education districts where NECT is active. The education districts are Bohlabela, Bojanala, Mount Frere, Uthungulu and Waterberg. Figure *1* is a map of the five study areas that were considered in this study.

This paper will discuss the assessment of the positional accuracy (an element of geographic information quality) of school location information. Geographic information quality elements can be used to discern if a dataset is fit for use for a specific application. According to International Standards Organisation (ISO)19113 Geographic Information – Quality principles, positional accuracy is one of the data quality components. For the purpose of this study, positional accuracy is defined as the closeness of agreement between the location (position) of a school in a dataset and the actual location (position) of a school on earth (reality).

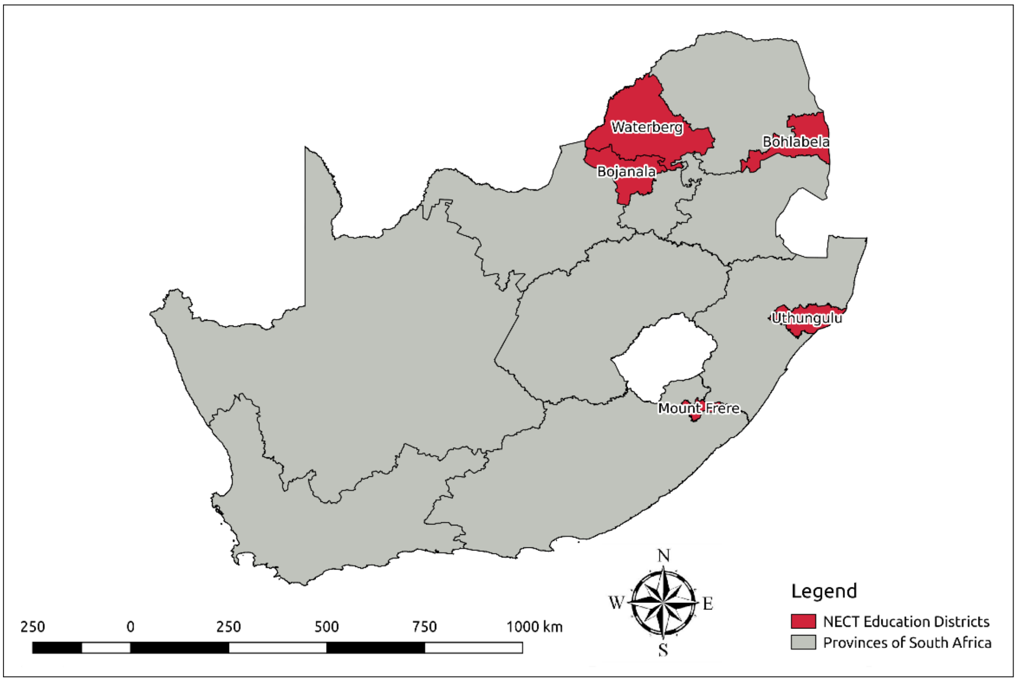

Figure 1. Study Areas – NECT Education Districts

3. Related work

Other authors have previously attempted to utilise South African school location data for various purposes. Previous efforts included the assessment and improvement of school location information in various study areas. Positional accuracy was an important element in all of the previous efforts.

The first effort to improve school location information occurred in 2011. The Chief Directorate: National Geo-Spatial Information (NGI) is known as South Africa's national mapping organisation and is a unit of the Department of Rural Development and Land Reform (DRDLR). In 2011 NGI undertook a task to improve the National EMIS dataset for schools in the Free State, Northern Cape and Western Cape (Sutherland, 2012). NGI noted that the location of each school was captured by teams using handheld Global Positioning System (GPS). NGI reported that the locations of schools were not accurately captured by the service providers that were contracted to populate the EMIS database (Sutherland, 2012). NGI attempted to improve the positional accuracy of the locations of the schools in the data by converting the coordinates of the school locations into Keyhole Markup Language (KML) format. NGI used Google Earth and their own imagery to identify the correct locations of schools. NGI used the name of each school, the number of pupils in a school and the street address of a school as supplemental information to provide guidance in identifying the correct location of a school (Sutherland, 2012). NGI collaborated with DBE and at the end of the effort, the locations of 3 989 schools were verified (Sutherland, 2012). Some of the schools in the three study areas are located in urban districts and some of the schools are located in rural districts. The type of area could have a big impact on being able to locate a school with services such as Google Maps, Google Earth, Google Street View, Bing Maps and HERE WeGo.

The second effort was led by Schmitz and Eksteen in 2014. The authors conducted a study to determine the effect of GIS data quality on school infrastructure planning in the City of Tshwane. The authors used the EMIS dataset (that contained positional errors) and a verified dataset (that did not contain positional errors) to determine locations of new schools. The authors noted that within the EMIS dataset, 50% of 501 schools contained positional errors and that no metadata was supplied with the EMIS dataset (Schmitz and Eksteen, 2014). In conclusion, the authors illustrated that by using data with poor positional accuracy and a lack of metadata, substantial amounts of money could be misspent through poor decision making (Schmitz and Eksteen, 2014). Ultimately, poor quality data will lead to poor decisions.

The third effort occurred in 2015, when the CSIR Meraka Institute conducted a review of existing school location datasets for a connectivity planning project for NECT in three educational districts namely Libode, Pinetown and Vhembe. CSIR utilised the 2013 EMIS dataset, the National Education Infrastructure Management System (NEIMS) dataset and the Statistics South Africa (StatsSA) Dwelling Frame dataset, together with Google satellite imagery for the project (Van Deventer et al., 2016). Every dataset (i.e. EMIS, NEIMS and Dwelling Frame) had its own table in a database. CSIR individually evaluated tables in a database to check for duplicate records (i.e. schools) based on NATEMIS, schools without coordinates, and to check if the educational district attribute of a school lies within the educational district border. CSIR then proceeded to do a comparison of the different tables with regards to coordinates, NATEMIS and school names. Schools that fell within a 100m distance of each other (with the same NATEMIS) were ignored for visual analysis purposes and schools that fell outside a 100m distance of each other were flagged for visual analysis. The visual analysis was conducted for all of these schools (Van Deventer et al., 2016). CSIR identified the following errors within the datasets. School names are spelt inconsistently between the datasets. The same school might have different coordinates between the datasets. There were schools without coordinates. Some of the school coordinates did not point to buildings at all (Van Deventer et al., 2016). Positional errors in the original data have an influence the quality of connectivity planning. Positional errors have an influence on radio propagation predictions for schools as well as distances to existing infrastructure calculations. These problems have a serious effect on infrastructure planning and implementation. This issue was specifically highlighted in the report by the Deputy Minister of Telecommunications and Postal Services to the Parliamentary Monitoring Committee in June 2017 (Parliamentary Monitoring Group Report, 2017). The SA-Connect project is currently more than two years behind target.

4. Methodology

The methodology for this study forms part of a larger initiative towards a process for planning and designing school connectivity in South Africa. In *Figure 2*, the high level methodology is highlighted.

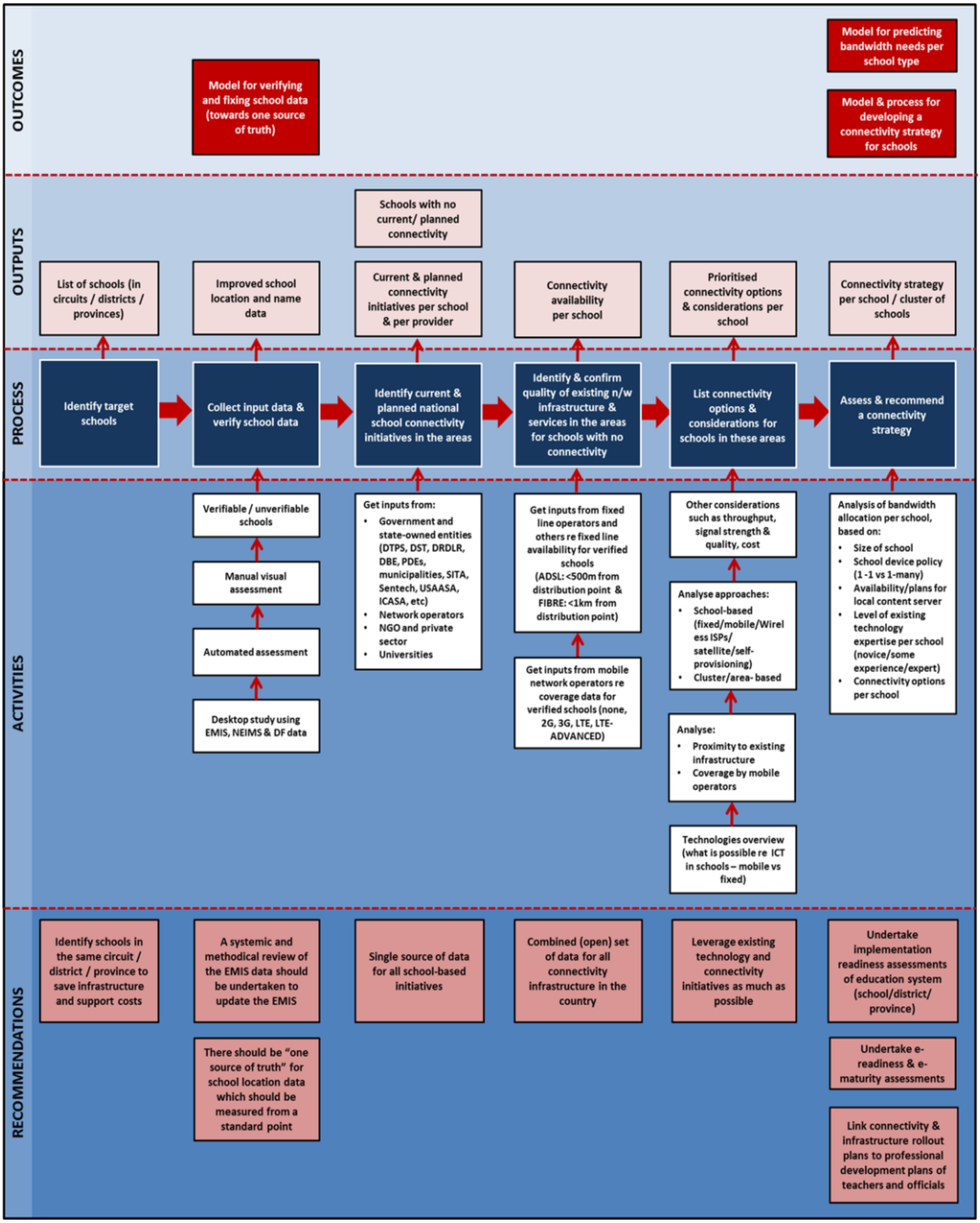
OUTCOMES
Model for verifying and fixing school data (towards one source of truth)
Model for predicting bandwidth needs per school type
Model & process for developing a connectivity strategy for schools
OUTPUTS
Schools with no current/ planned connectivity
List of schools (in circuits / districts / provinces)
Improved school location and name data
Current & planned connectivity initiatives per school & per provider
Connectivity availability per school
Prioritised connectivity options & considerations per school
Connectivity strategy per school / cluster of schools
PROCESS
Identify target schools
Collect input data & verify school data
Identify current & planned national school connectivity initiatives in the areas
Identify & confirm quality of existing n/w infrastructure & services in the areas for schools with no connectivity
List connectivity options & considerations for schools in these areas
Assess & recommend a connectivity strategy
ACTIVITIES
Verifiable / unverifiable schools
Manual visual assessment
Automated assessment
Desktop study using EMIS, NEIMS & DF data
Get inputs from:
• Government and state-owned entities (DTPS, DST, DRDLR, DBE, PDEs, municipalities, SITA, Sentech, USAASA, ICASA, etc)
• Network operators
• NGO and private sector
• Universities
Get inputs from fixed line operators and others re fixed line availability for verified schools (ADSL: <500m from distribution point & FIBRE: <1km from distribution point)
Get inputs from mobile network operators re coverage data for verified schools (none, 2G, 3G, LTE, LTE-ADVANCED)
Other considerations such as throughput, signal strength & quality, cost
Analyse approaches:
• School-based (fixed/mobile/Wireless ISPs/ satellite/self-provisioning)
• Cluster/area- based
Analyse:
• Proximity to existing infrastructure
• Coverage by mobile operators
Technologies overview (what is possible re ICT in schools – mobile vs fixed)
Analysis of bandwidth allocation per school, based on:
• Size of school
• School device policy (1 -1 vs 1-many)
• Availability/plans for local content server
• Level of existing technology expertise per school (novice/some experience/expert)
• Connectivity options per school
RECOMMENDATIONS
Identify schools in the same circuit / district / province to save infrastructure and support costs
A systemic and methodical review of the EMIS data should be undertaken to update the EMIS
There should be "one source of truth" for school location data which should be measured from a standard point
Single source of data for all school-based initiatives
Combined (open) set of data for all connectivity infrastructure in the country
Leverage existing technology and connectivity initiatives as much as possible
Undertake implementation readiness assessments of education system (school/district/ province)
Undertake e-readiness & e-maturity assessments
Link connectivity & infrastructure rollout plans to professional development plans of teachers and officials

Figure 2. High-level Methodology

4.1 Data

4.1.1 EMIS

The DBE serves as the custodian of the EMIS dataset that is publicly available. Provincial departments are responsible for the maintenance of the dataset. The data contained in the dataset is collected from SNAP surveys and annual school surveys conducted at ordinary schools, special needs schools, Adult Basic Education Training (ABET) centres and Further Education Training (FET) colleges (van Wyk, 2015). Only ordinary schools were considered for this study. The "Quarter 4 of 2015: March 2016" dataset was used for this project as it was the most up to date version of the dataset available at the time of this study (Department of Basic Education, 2016).

The dataset contains several attribute values for each record (i.e. school) but only the NATEMIS (unique identifier), longitude and latitude attribute values were considered for the purpose of this study.

4.1.2 NEIMS

The NEIMS dataset is compiled from data collected during audits at public schools (quarterly) by provincial education departments in South Africa. The information was obtained from the DBE. The NEIMS III dataset was used for this project (Moloto, 2016).

4.2 Methods

A manual process was followed as it is not possible to fully automate the detection of school buildings in satellite imagery (by using machine learning). It is not always the case in South Africa, that a school building is U-shaped or L-shaped with a sports field. Some school buildings have the shape of houses and some schools are located in small centres. The methodology followed in this project is discussed below.

4.2.1 Data Pre-Processing and Import

The datasets were converted to GIS formats as the datasets were provided in Microsoft Excel Spreadsheet (XLS) and Garmin MapSource Database (GDB) formats. Both datasets were clipped to the boundaries of the five education districts. Some of the schools included in the EMIS dataset, did not have any coordinates and a geocoding process was performed to try to assign coordinates to these schools based on their name or address. If coordinates were assigned to a school and the school was contained in one of the five educational districts, the school was added to the dataset. For each education district boundary, the EMIS and NEIMS datasets were combined into a single shapefile. The combined shapefiles were imported into a database with one table per education district.

4.2.3 Data Analysis

A search was conducted to identify duplicates within the EMIS and NEIMS datasets based on the NATEMIS (i.e. schools that appear more than once in the same dataset). A second search was conducted to identify duplicates between the EMIS and NEIMS datasets (i.e. schools that appear in both datasets). A visual inspection of the duplicate entries was then conducted to find the record that was closest to the actual school by visual verification.

4.2.4 Visual Inspection and Desktop Study

Bing Maps Aerial and Google Satellite imagery were used to find the school locations by visual inspection. For duplicate records, the record with the location that was found to be closest to the actual school location was moved to the actual school location if it was not correct in the initial dataset. Unique records were moved to the actual school location if it was not correct in the initial dataset. Where a school was not located by visual inspection an online search was conducted to find information on the school's location and Google Street View was used to identify the location, if necessary and possible (due to coverage). As the records were processed they were marked as checked and if the record was a duplicate it was marked as duplicate in the appropriate database field to be deleted later on. After all records were processed, the records that were marked as duplicate were deleted from the dataset.

4.2.5 Data Visualisation, Export and Additional Processing

Maps were created based on the original and updated school locations in the combined EMIS-NEIMS dataset. The new dataset was exported from the database in Comma Separated Value (CSV) format and converted to shapefile format for use in a GIS. Based on the new combined EMIS-NEIMS dataset, reverse geocoding techniques were applied to identify addresses for the school locations.

5. Results and discussion

Table 1 contains the results of a comparison done between the original location of a school that appears in the EMIS dataset and in the NEIMS dataset (i.e. NATEMIS appears in the EMIS dataset and the same NATEMIS appears in the NEIMS dataset). *Table 1* also contains statistics on the amount of schools with a NATEMIS that only appears in one dataset (i.e. Not in Dataset).

Table 1. Difference in Distance between EMIS and NEIMS Schools Locations

	Uncertain/Unverified		**Verified**	
Difference in Distance	Number of Schools	Percentage of Schools	Number of Schools	Percentage of Schools
0m	55	2.17%	492	19.45%
>0m & <=100m	25	0.99%	795	31.42%
>100m & <=1000m	25	0.99%	409	16.17%
>1000m	50	1.98%	130	5.14%
Not in Dataset	117	4.62%	432	17.08%
Total	**272**	**10.75%**	**2258**	**89.25%**

The positional accuracy required from the school locations in the EMIS dataset and the NEIMS dataset will vary depending on the type of application. As illustrated in *Table 1* and *Figure 3*, one can observe that out of 2530 schools, 549 schools do not appear in both datasets (i.e. only in the EMIS dataset **OR** only in the NEIMS dataset). That amounts to approximately 22% of schools in the study areas that only have one location. The remainder of the schools, approximately 78% appear in both of the datasets (i.e. in the EMIS dataset **AND** the NEIMS dataset) and therefore the locations of those were compared.

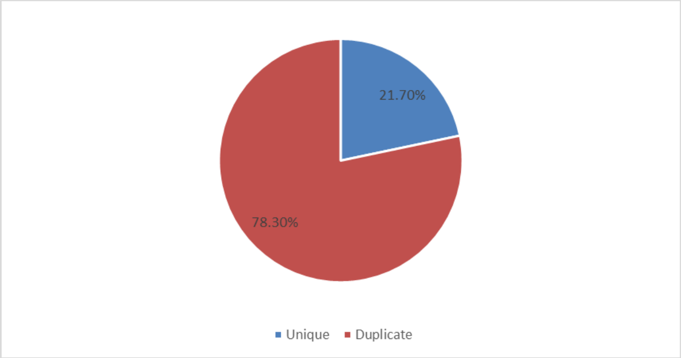

Figure 3. Percentage of Unique and Duplicate Schools

Out of 2530 schools about 89% of schools could be verified through visual inspection and desktop study. *Figure 4* is an illustration of a school with a verified location. The remaining 11% of schools could not be verified because of a number of reasons such as no image coverage in the area (and imagery might be outdated in certain areas), cloud cover in imagery, no Google Street View coverage, no address data available in gazetteers and no additional information available on the internet. *Figure 5* is an illustration of a school with a without a verified location.

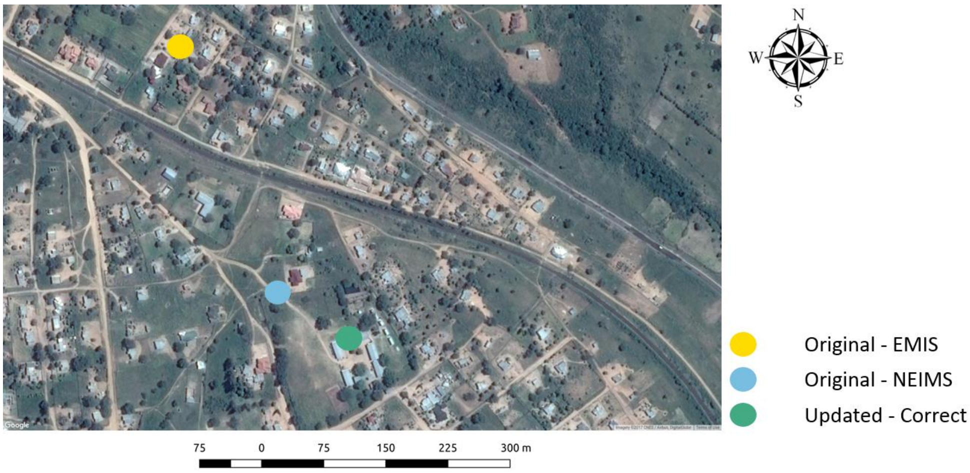

Figure 4. Verified School – NATEMIS 800034916 Original Locations and Updated Locations

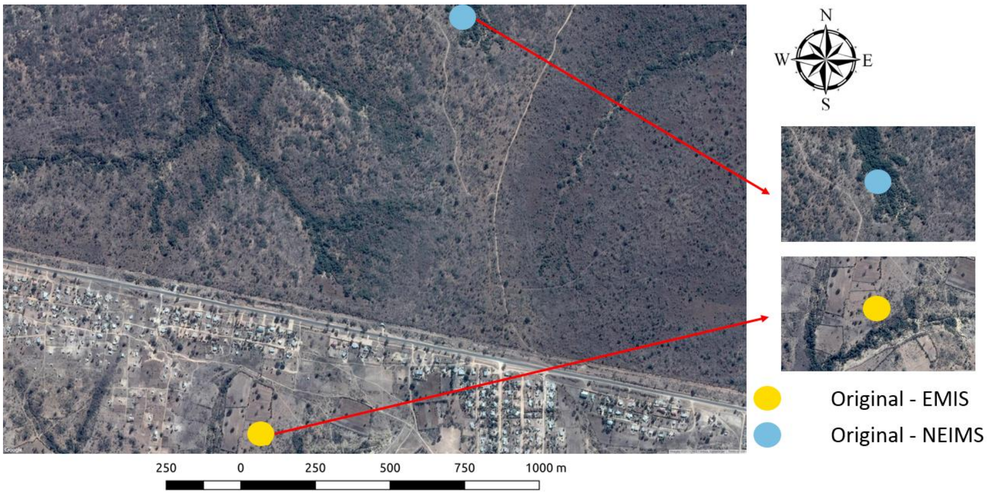

Figure 5. Unverified School – NATEMIS 800034930 Original Locations

School locations that appeared in only one dataset were more difficult to verify because two locations provided a better indication of the possible area where the school might be located. It should be noted that the study areas are rural districts. Rossouw and Kgope (2007) stated that most households without formal addresses in South Africa are in rural areas and informal settlements. Therefore, the geocoding and reverse geocoding techniques that were applied were unsuccessful.

Out of about 89% school locations that could be verified with visual inspection and desktop study, approximately 81% of schools appeared in both datasets and only approximately 19% of schools only appeared in one dataset.

Out of the 2258 verified school locations (see *Figure 6*), around 27% of schools that appear in both datasets have a distance of 0m between the locations in the respective datasets. The majority of schools (approximately 44%) that were verified and appear in both datasets had a distance of greater than 0m and smaller than 100m between the locations in the respective datasets. The remainder of the schools that were verified and appear in both datasets, had a distance of greater than 100m and smaller than 1km (approximately 22%), or a distance of greater than 1km (approximately 7%) between the locations in the respective datasets.

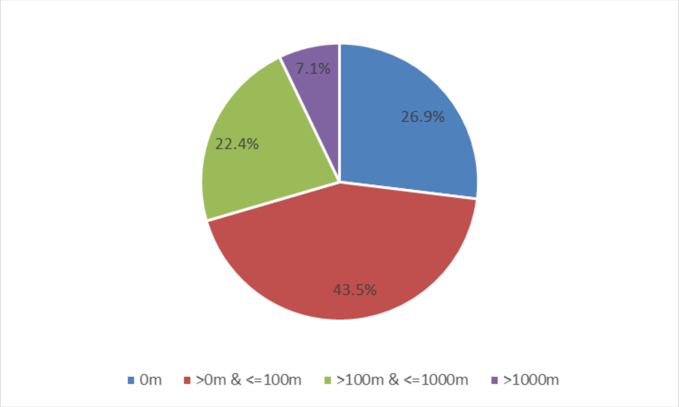

Figure 6. Difference in Distance between EMIS and NEIMS Schools Locations

Table 2 contains the results of a comparison done between the location of a school that appears in the EMIS dataset and in the new updated dataset. *Figure 7* shows the difference in distance for verified schools only. Only around 2% of the verified schools were not moved. Approximately 57% of the verified schools were moved less than 100m from the location in the original dataset. About 19% of verified schools were moved between 100m and 1km; and about 5% of schools were moved more than 1km from the original location in the EMIS dataset.

Table 2. Difference in Distance between EMIS and New Dataset Schools Locations

	Uncertain/Unverified		**Verified**	
Difference in Distance	Number of Schools	Percentage of Schools	Number of Schools	Percentage of Schools
0m	71	2.81%	48	1.90%
>0m & <=100m	48	1.90%	1278	50.51%
>100m & <=1000m	30	1.19%	426	16.84%
>1000m	22	0.87%	108	4.27%
Not in Dataset	101	3.99%	398	15.73%
Total	**272**	**10.75%**	**2258**	**89.25%**

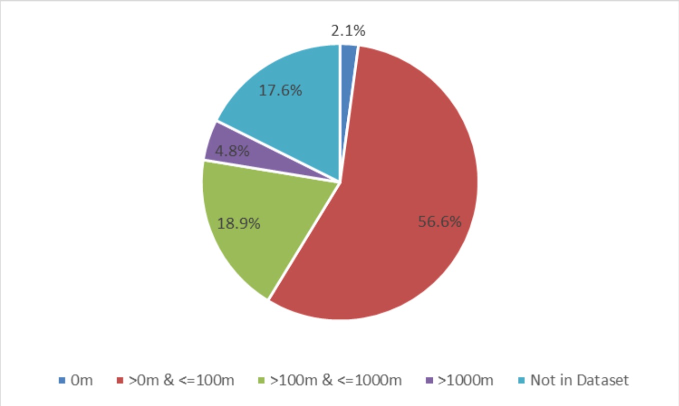

Figure 7. Difference in Distance between EMIS and New Dataset Schools Locations

Table 3 contains the results of a comparison done between the location of a school that appears in the NEIMS dataset and in the new updated dataset for verified and unverified data. *Figure 8* shows the difference in distance for verified schools only. Approximately 18% of the verified schools were not moved. Approximately 76% of the verified schools were moved less then 100m from the location in the original dataset. About 3% of verified schools were moved between 100m and 1km; and approximately 1% of schools were moved more than 1km from the original location in the NEIMS dataset.

Table 3. Difference in Distance between NEIMS and New Dataset Schools Locations

	Uncertain/Unverified		**Verified**	
Difference in Distance	Number of Schools	Percentage of Schools	Number of Schools	Percentage of Schools
0m	114	4.51%	407	16.09%
>0m & <=100m	82	3.24%	1725	68.18%
>100m & <=1000m	25	0.99%	70	2.77%
>1000m	35	1.38%	22	0.87%
Not in Dataset	16	0.63%	34	1.34%
Total	**272**	**10.75%**	**2258**	**89.25%**

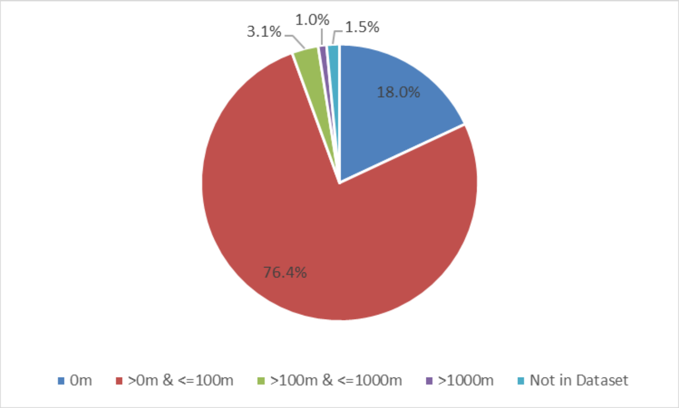

Figure 8. Difference in Distance between NEIMS and New Dataset Schools Locations

The results depicted in *Figure 7* and *Figure 8*, indicate that the positional accuracy of the school locations in the NEIMS dataset is better than the positional accuracy of the school locations in the EMIS dataset. As DBE is the custodian of both of the abovementioned datasets, it is clear that there exist some duplication of effort in the department as the two datasets contain different school locations. Not only do the datasets contain different locations for most of the schools, but the locations in the five study areas also contain positional errors.

Costs are incurred during the process of location surveying as individuals have to travel to capture coordinates with GPS. It would be more cost efficient if effort is not duplicated in future. The South African Spatial Data Infrastructure (SASDI) was established by the Spatial Data Infrastructure Act 54 of 2003. SASDI is the "technical, institutional and policy framework to facilitate the capture, management, maintenance, integration, distribution and use of spatial information". One of the objectives of SASDI is to eliminate duplication in the capturing of spatial information. This means according to an Act of Parliament, duplication in data capturing should be avoided.

According to the Spatial Data Infrastructure Act 54 of 2003, a data custodian should also ensure that metadata is available. No metadata was supplied with the EMIS dataset or the NEIMS dataset. Fourie (2015) argued that data custodians need to capture metadata because it will include some information on the data to determine its fitness of use for a specific application. It is important that DBE considers supplying metadata with their datasets.

The non-delivery of textbooks in South Africa is a common issue. Chrisholm (2013) cited one of the reasons that compromised the delivery of textbooks in Limpopo between 2011 and 2012 being that the school location information was inaccurate. It is imperative that DBE improve the positional accuracy of the school locations (nationally) as a step in order to mitigate problems related to the non-delivery of textbooks. It was observed that the March 2016 dataset contained outdated data (e.g. from 2007 and 2010).

Study areas (i.e. education districts) were mentioned at the start of the paper, but the results were presented as a whole to provide an overall view of positional errors in the school location information. The aim of this paper was by no means to criticise the DBE or single out the work of specific districts, but to provide a general overview of the errors and illustrate the importance of why their only needs to be one dataset with good positional accuracy.

6. Conclusion

In order to do effective planning and decision making in a variety of sectors, accurate school location information is required. This study compared two available school location datasets to investigate the accuracy of school locations in the datasets. Many of the school locations, when inspected visually, were found to be inaccurate. The school locations were corrected and compiled into a new dataset. An analysis was conducted to determine the positional accuracy of the coordinates in the EMIS and NEIMS datasets compared to the new dataset. The NEIMS dataset proved to have better positional accuracy, but still contained positional errors. Some of the school locations could not be verified and therefore CSIR recommends field work by well-trained individuals. DBE should consider to recapture all of the school locations in South Africa **once** in future (by contracting GPS trained individuals or crowdsourcing by school principals), do quality assessments on the data as it is captured, and consolidate the data into a single dataset and update and maintain it. If the DBE manages to improve the positional accuracy of school location information in South Africa, decisions that are made based on this data will have better quality outcomes, costs will be significantly decreased and some service delivery issues related to education will be mitigated.

7. References

Chrisholm, L 2013. Understanding the Limpopo textbook saga, viewed 12 July 2017, <http://www.hsrc.ac.za/en/review/hsrc-review-september-2013/understanding-the-limpopo-textbook-saga>.

Department of Basic Education 2014, National Guidelines on: Completing the SNAP Survey for Ordinary Schools, viewed 16 August 2016, <http://www.education.gov.za/Portals/0/Documents/Publications/National%20Guidelines%20on%20SNAP%20Ordinary%202014.pdf?ver=2015-01-29-160331-687>.

Department of Basic Education 2016, EMIS - Education Management Information Systems, viewed 16 August 2016, <http://www.education.gov.za/Programmes/EMIS.aspx>.

Department of Basic Education 2016, Schools Masterlist Data, viewed 16 August 2016, <http://www.education.gov.za/Programmes/EMIS/EMISDownloads.aspx>.

Department of Telecommunications and Postal Services, 2015, SOUTH AFRICA CONNECT: CREATING OPPORTUNITIES, ENSURING INCLUSION South Africa's Broadband Policy, 20 NOVEMBER 2013, <http://www.dtps.gov.za/index.php?option=com_phocadownload&view=category&id=21&Itemid=333>.

Fourie, H, 2015, The need for data custodians to capture meaningful metadata , in Proceedings of the 2015 Geomatics Indaba, Ekurhuleni, 11-13 August.

Malebye, J (DBE) 2015, Education District Shapefile.

Moloto, C (DBE) 2016, NEIMS – National Education Infrastructure Management System.

National Education Collaboration Trust 2017, About the NECT, viewed 12 July 2017, <http://nect.org.za/about/about >.

Parliamentary Monitoring Group Report, 2017, SA Connect: Department progress report, with Deputy Minister, viewed 1 July 2017, <https://pmg.org.za/committee-meeting/24580/>.

Republic of South Africa 2003, Spatial Data Infrastructure Act No.54 of 2003, Government Gazette No. 25973 of 2004.

Republic of South Africa 2012, National Development Plan 2030: Our Future – make it work.

Rossouw, P and Kgope, K 2007, Rural Addressing in South Africa, PositionIT, viewed 12 July 2017, <http://www.ee.co.za/wp-content/uploads/legacy/PositionITSept-Oct%2007-p66-71.pdf>.

Schmitz, P and Eksteen, S 2014, The Effect of GIS Data Quality on Infrastructure Planning: School Accessibility in the City of Tshwane, South Africa, in Proceedings of the Second AfricaGEO Conference, Cape Town, 1-3 July.

Sutherland, D 2012, Putting schools on the Map, PositionIT, viewed 16 August 2016, <http://www.ee.co.za/article/ngi-011-06.html>.

Van Deventer, A., Roux, K., Ford, M. and Smith, R 2016, Review and recommendations on South African school location data for National Education Collaboration Trust, CSIR Report.

Van Wyk, C 2015, An overview of education data in South Africa: an inventory approach, Stellenbosch Economic Working Papers: 19/15, viewed 16 August 2016, <www.ekon.sun.ac.za/wpapers/2015/wp192015/wp-19-2015.pdf>.