

GROUNDWATER MOBILE APP DEVELOPMENT TO ENGAGE CITIZEN SCIENCE

STEFANUS RAINIER DENNIS & INGRID DENNIS

Centre for Water Sciences and Management, North-West University, South Africa

ABSTRACT

Public domain borehole information in South Africa is generally stored in the National Groundwater Archive (NGA) and the Groundwater Resources Information Project (GRIP) databases of which both are centralized databases. The GRIP database is updated by the Department of Water and Sanitation, but only covers one of the nine provinces. The NGA on the other hand covers the entire country, however there is a backlog of borehole information that needs to be captured, and it has limited time series data. The reason for the poor time series data is twofold: (i) groundwater monitoring is expensive due to the distributed nature of the resource (the NGA consists of approximately 280,000 boreholes over an area of 1,225,986 km²); and (ii) consultants tend not to upload data to the national databases as the data is seen as a competitive advantage. During the recent drought experienced in the Western Cape Province (2015–2018), citizens of local communities took to social media, reporting on rainfall and groundwater levels within their communities. With the dams drying up people started targeting groundwater with approximately 30,000 boreholes being drilled. This led to the development of a mobile app for both citizens and groundwater professionals. This app allows logging of borehole information via smart phones. One of the main challenges with populating databases is the verification of data. The mobile app introduces a type of block chain approach where all data is accepted, but marked as low confidence until verified by a trusted user. The vision for the app is a “live” hydro census and even if only water levels are captured, it would improve groundwater management by applying data mining techniques for trend analysis in future.

Keywords: citizen science, borehole database, mobile app.

1 INTRODUCTION

Public domain borehole information in South Africa is generally stored in the National Groundwater Archive (NGA) and the Groundwater Resources Information Project (GRIP) database of which both are centralized databases. The Limpopo GRIP database is regularly updated by means of a term contract set up by the Department of Water and Sanitation (DWS). The NGA is known for a backlog of borehole information that still needs to be captured and the GRIP database should ultimately also form part of this database.

The NGA database relies on users to upload captured borehole information to the database, but this is not happening for various reasons. The flow of data for the current situation is depicted in Fig. 1, where the partially filled arrows refer to the partial transfer of data. The partial transfer of data between the NGA and GRIP databases relates to the reluctance of users to transfer data for the following reasons:

- The GRIP users and administrators want to keep the GRIP database separate and don't want merging of data with the NGA due to the fact that they perceive the NGA data to be substandard.
- Third parties are reluctant to upload data, due to the fact that data has monetary value and if a person has a dataset which someone else does not, it is to that person's advantage.

The partial data transfer between the NGA and the users is mainly due to system and institutional constraints from the DWS. System constraints refer to the capability of the



system to automatically send all query results to the users. Chemistry related queries are performed by a person rather than the system and users don't always get a response when querying chemistry data. Institutional constraints refer to backlogs that exist due to not having enough capacity to capture data. The combined borehole distribution of the NGA and GRIP databases is presented in Fig. 2.

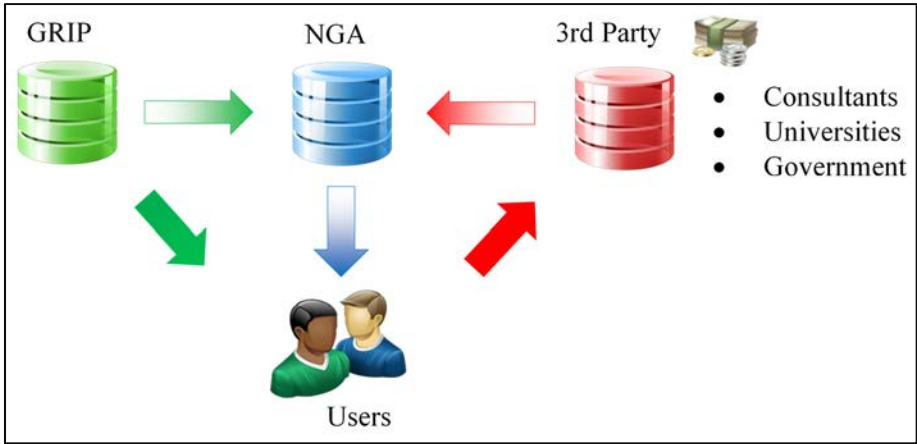


Figure 1: Current data flow between entities.

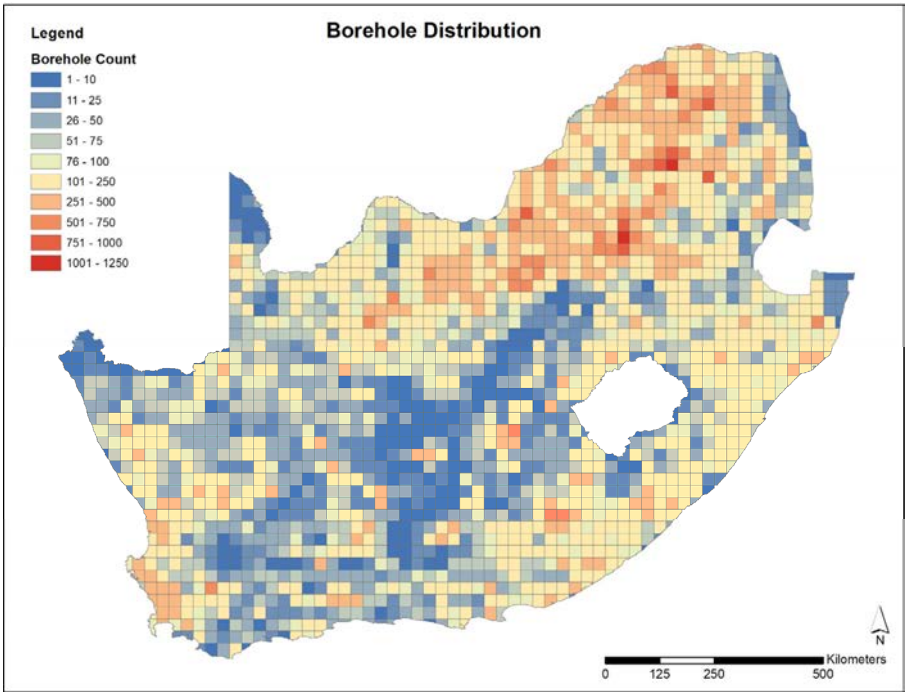


Figure 2: Borehole distribution of the NGA and GRIP databases [9].



2 CITIZEN SCIENCE

Citizen science is defined by the Oxford dictionary as the collection and analysis of data relating to the natural world by members of the general public, typically as part of a collaborative project with professional scientists. Other aliases for “citizen science” include “amateur science”, “crowd sourced science”, “volunteer monitoring” and “public participation in scientific research” [1].

2.1 Background

Some of the first citizen science projects started with bird counts in the early twentieth century [2]. The phrase “citizen science” itself was coined in the mid-1990s. Irwin, a sociologist, defined it both as “science which assists the needs and concerns of citizens” and as “a form of science developed and enacted by the citizens themselves” [2]. From Irwin’s definitions it is clear that the citizens should have vested interest in the relevant project.

2.2 The smartphone era

With the advent of mobile communication technologies such as smartphones, citizen science has become increasingly popular. Apple opened the door to this disruptive technology at the time in 2007, with the launch of the iPhone. Consumers and scientists alike, embraced this new technology and today a world without smartphones is unimaginable. The worlds of scientists and citizen science are now converging, with scientists harnessing the power of citizen science with newer and smarter technologies [3].

The fact that a smartphone can take a photo and associate a global positioning system (GPS) coordinate and a time stamp with it allows for the collection of both spatial and temporal data, which prove invaluable in many scientific fields especially in the environmental sciences. Current day smartphones have an array of onboard sensors that can be used in various applications. Newer Japanese models are even fitted with radiation detection sensor [3]. Scientists are continuously developing new techniques to harness the power of the onboard sensors.

Concerns regarding the use of technology in citizen science projects have also been raised. Not all groups have access to this type of technology and therefore are excluded. The steadily increasing number of smartphone apps can also prove overwhelming for users unfamiliar with these tools [4].

2.3 Suitability and success factors

Both the suitability of the project and success factors should be considered when embarking on a citizen science project. Six broad areas to review the suitability [5] is presented in Fig. 3 and the associated questions to be answered for each of these areas is given in Table 1.

In addition to the six broad areas defined a decision framework [5] exists to evaluate each considered project in terms of its potential to succeed as a citizen science project. The decision framework considers basic questions like online surveys, safety in making observations, are sensors required and the spatial scale of the project. Once these questions have been answered, the decision framework categorizes the project based on spatial and temporal requirements. Lastly, for each type of category, questions regarding access to sites, site types and protocols result in the final recommendation regarding potential of the project.

Six success factors for citizen science were identified by Pocock et al. [5] and are depicted in Fig. 4. Although the aforementioned work was focused on coastal and ocean research, these principles are generic and applicable to all citizen science efforts. Most of these factors are interrelated at some level as illustrated in the short discussion that follows.





Figure 3: Suitability criteria for a citizen science project [5].

Table 1: Questions asked with respect to suitability criteria [5].

Suitability criteria	Typical question to be answered
Clarity of the aim	Do you have a precise and clearly-defined aim for your citizen science project?
Importance of engagement	Can you extend your engagement activity into meaningful and relevant citizen science or should you simply undertake excellent engagement for its own sake?
Resources available	Do you have sufficient resources available to ensure you can support your volunteers for the entirety of the project?
Scale of sampling	Do you need many people (or volunteer time or commitment) to achieve your aims?
Complexity of the protocol	Is your protocol practical for volunteer involvement? Are you expecting too much from the volunteers?
Motivations of participants	Does your project resonate with potential volunteers, and are there clear and appropriate triggers for people to make records?



Figure 4: Factors of success in citizen science [4].

First and foremost, clear goals should be set for the project. Keep the project focussed and avoid the “big bang” approach, where it is attempted to collect all possible types of data just because it can be done. This in turn relates to the engagement of citizens – if the collection process is long and tedious, people will lose interest.

Reliable data is a prerequisite for an improved database and an improved database will contribute to the science through proper analysis of the acquired data. Data quality and verification underpin reliable data and most data quality problems can be addressed if they are properly anticipated. The simplest measure is training and communication to the network of people participating in the data collection.

2.4 Growth in citizen science

Significant growth of the citizen science projects has been recorded over the last decade (Fig. 5) as indicated by the SciStarter [1] repository, which keeps track of citizen science projects. The field of citizen science is largely decentralized, which makes the tracking of these projects difficult [2]. Some academics fear that the public is getting fatigued by all the options, and noted that participation in some projects has declined [2].

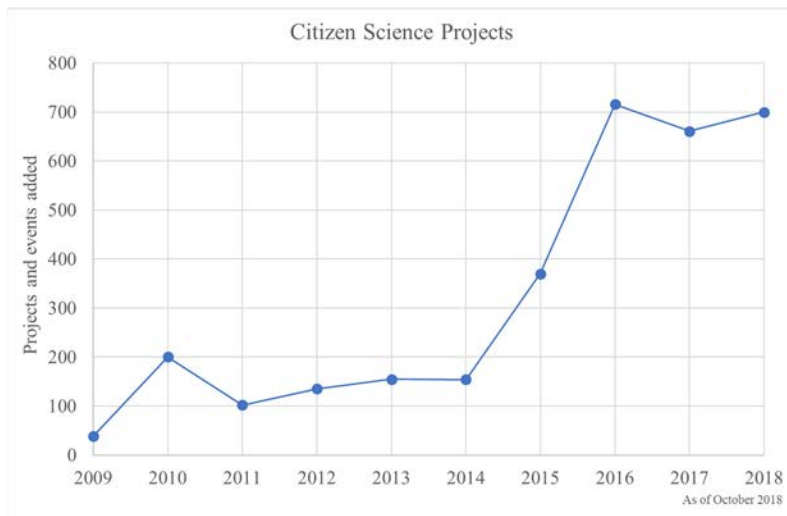


Figure 5: Factors of success in citizen science. (Source: Adapted from Garcia-Soto et al. [4].)

3 HUMAN BEHAVIOUR DURING A WATER CRISES

3.1 Setting the scene

In 2018, the City of Cape Town experienced its worst drought in over a century. The historic dam levels from 2016 are presented in Fig. 6. The water crisis peaked during mid-2017 to mid-2018 where water levels hovered between 15 to 30% of total dam capacity. In late 2017, the first mentions of plans for “Day Zero” were made (indicated as the estimated cut-off level in Fig. 6), a shorthand reference for the day when the demand for water to meet essential needs exceeds the supply [6].

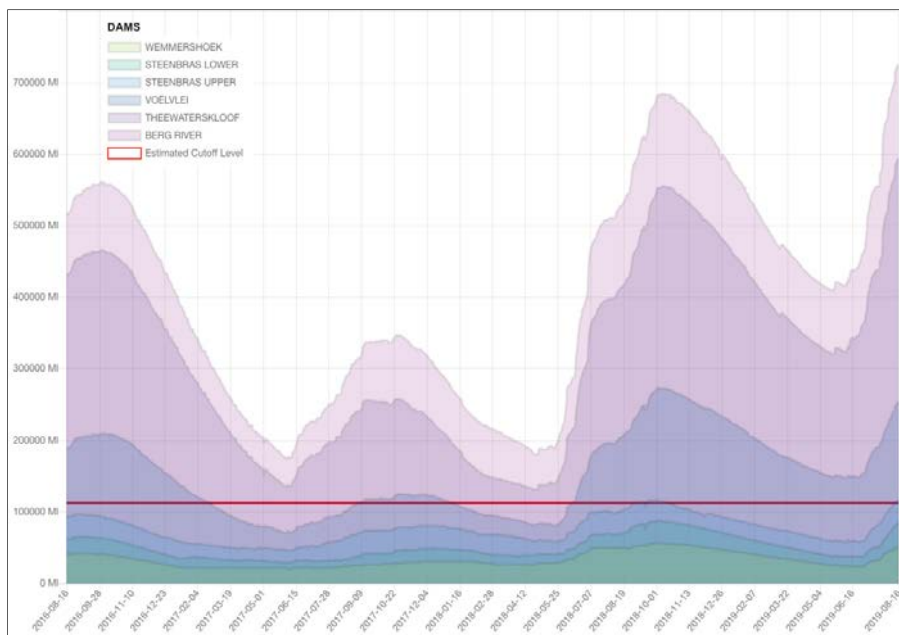


Figure 6: Historical dam levels for the Cape Town region in South Africa [8].

“Day Zero” would herald the start of Level 7 water restrictions, where municipal water supplies would largely be switched off and residents would have to queue for their daily ration of water, making the City of Cape Town the first major city in the world to potentially run out of water [7].

3.2 Response to water crises

During the ongoing drought people resorted to groundwater as a source of water. This resulted in an estimated 30,000 boreholes being drilled in and around the Cape Town. In general, a water use license is required for water use in South Africa under the National Water Act (No. 36 of 1998). However, the water law makes provision for water-use that is termed “Schedule 1” use. Schedule 1 water uses are generally low-volume, low-impact activities that are consistent with domestic use, livestock watering, recreational use and the use of water for emergencies.

This water use is permissible and does not require licensing or registration. The concern from a water management point of view is the question: Is 30,000 boreholes in close proximity to each other still a low-impact activity? A further concern is that since no licensing and registration is required for Schedule 1 water-use, the positions of the majority of these boreholes will not be known, unless the drilling contractors and consultants provide this information to the authorities.

3.3 Water status reporting on social media platforms

As people were affected more and more by the water shortage, they took to social media platforms and started reporting rainfall and to a lesser extent water levels in the areas where

they lived. The data resolution obtained in certain areas were astonishing – people had a vested interest to report on the status of water as this could assist in management.

The problem with social media platforms is that the data is highly distributed and is cumbersome to collate for management purposes. This response from the public sparked the development of a borehole mobile app to assist with borehole identification and monitoring to enable better water management in future. The following quote from John Thorson never held truer than during the time of drought:

“Water links us to our neighbour in a way more profound and complex than any other.”

4 BOREHOLE DATA CHALLENGES

Some inherent challenges exist with respect to the existing borehole data and the borehole entity itself when it comes to measurements.

4.1 Current status of national borehole databases

The summary statistics of the national borehole databases is presented in Table 2 and the spatial distribution is shown in Fig. 2. It should be noted that the parameters presented in Table 2 is an abbreviated list of parameters available in these databases. The following observations are made from Table 2:

- 46% of boreholes have one or more water level;
- 14% of boreholes have one or more electrical conductivity (EC) measurements;
- 6% of boreholes have one or more chemistry datasets;
- 28% of boreholes have yield values;
- 38% of boreholes have water strike data;
- 50% of boreholes have borehole logs.

Table 2: Summary statistics of national borehole databases [9].

Count	Description	NGA	GRIP
Borehole	Total borehole count in each database. Roughly 11% of the GRIP boreholes are also present in the NGA database.	276,333	26,912
Water level	The number of boreholes that have one or more water levels associated with it.	129,615	10,333
Electrical conductivity	The number of boreholes that have one or more EC measurements	33,153	7,884
Chemistry (major anions and cations)	The number of boreholes that have one or more chemistry analysis associated with it (2008 data as no new chemistry data could be obtained).	12,173	7,050
Yield	The number of boreholes that have one or more yield values associated with water strikes.	75,327	10,002
Water strikes	The number of boreholes with one or more water strikes associated with it.	112,880	2,386
Borehole logs	Total number of boreholes with a borehole log.	147,837	3,403

It is clear that these databases are sparsely populated and have poor temporal data (not presented here). Monitoring programmes are expensive to operate and the majority of boreholes are located in rural areas, as these are used for water supply where no surface water infrastructure exists.

The DWS in South Africa does have a dedicated borehole monitoring network that collects long-term borehole data, but these monitoring points only make up a fraction of the total number of boreholes and were chosen as “representative” points of various aquifers in South Africa. Groundwater data is however highly variable over short distances, making the selection of representative points an impossible task.

4.2 Borehole localities

Although military development of the GPS started in the late 1950s, GPS technology only became available to the public in the 1990s. The borehole positions captured in the NGA before GPS availability was done by assigning the borehole to the relevant farm or plot centroid. The result of this, is that a few of the 276,333 records will be located in the field as the associated coordinate is the farm or plot centroid. It is therefore not uncommon to obtain multiple boreholes assigned to the same position (Fig. 7), with a small offset to prevent duplicate coordinate positions in the database.

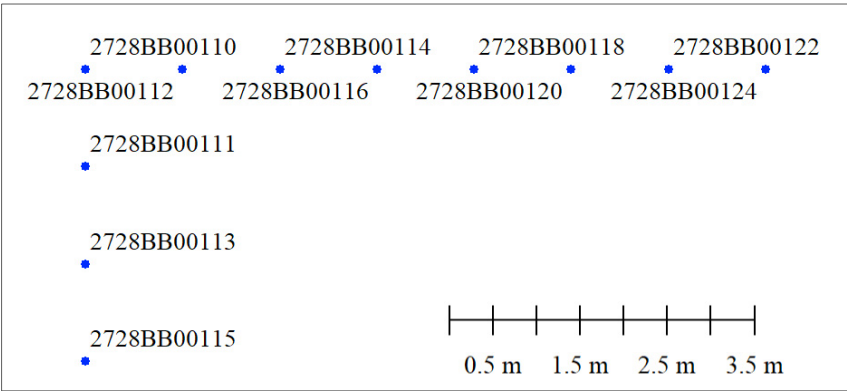


Figure 7: Boreholes with offset to prevent duplicate coordinates.

Unfortunately, the fact of the matter is that some of the boreholes present in the NGA, will never be located. This does not mean that these boreholes should be removed, as associated data like geological logs do provide information on the surrounding area even though the exact position is unknown.

4.3 Borehole parameters

From a citizen science point of view, to identify a borehole position and take a photo is elementary. However, borehole measurements require specific equipment and training. The solution to this challenge is to present the public with selection lists to give an indication of the status of the borehole if they are not equipped to perform proper measurements. The critical data captured remains the borehole locality, so that the position is known and the site can be revisited in future to carry out proper measurements.

5 METHODOLOGY

The methodology consists of evaluation of the project in terms of its suitability as a citizen science project and the selection of an appropriate platform for the development.

5.1 Evaluate the suitability and potential

The project potential as citizen science project, was evaluated by applying the suitability criteria and decision framework described by Pocock et al. [5]. The aforementioned is no guarantee to success of the final project, but merely provides guidance to success factors to be considered based on lessons learnt from other citizen science projects.

5.2 Software solution

The decision between custom software development and an off-the-shelf solution, is one that could be detrimental to the success of the project. Both these approaches have pros and cons to be considered. Table 3 provides a comparison of the key features to be considered in the context of the aforementioned solution types.

Table 3: Comparison of custom development vs. off-the-shelf software solution.

Feature	Custom development	Off-the-shelf solution
Cost and time	Generally, cost more and development time is longer.	Cheaper in comparison to new software development, but at the cost of functionality.
Maintenance	Done by developer or 3rd party if source code is available.	Maintenance generally done by seller and included in licensing.
Solution maturity	New software is subjected to testing and teething problems with new developments is common.	Generally, a stable solution as it has already matured over time.
Operating system compatibility	As operating systems change on target devices, the developer or 3rd party needs to ensure compatibility. This could also be part of product maintenance and updates.	From a commercial product point of view, this will form part of the product life cycle and likely be included in the maintenance.

When considering the comparison in Table 3, the ESRI App Studio® for ArcGIS® is consider a favourable solution, for the following reasons:

- GIS is ESRI's core business and App Studio® allows mobile development that integrate directly with existing data and maps.
- App Studio® provides the ability to build a 100% custom app, use templates or use existing apps e.g. Survey123 to be modified according to the developer's specification by the use of an extensive library with tested source code, which dramatically reduces development time.
- App Studio® further allows the deployment of the app to multiple mobile (and desktop) operating systems, making use of a single code base.

The proposed development platform has obvious advantages that relate to the success of the project, but the six success factors presented by Garcia-Soto et al. [4] should form the cornerstone of the design and implementation as far as possible.



5.3 Online database







The mobile app will require an online database to present users with existing borehole positions and related data. The ESRI solution already provides this functionality and facilitates seamless integration with the App Studio® solution. Both the GRIP and NGA databases form the basis of the online borehole database as these databases, even though poorly populated in certain respects, form a good basis to start from.

6 RESULTS AND DISCUSSION

6.1 Suitability and project potential

The suitability of the project for citizen science was evaluated making use of the criteria described by Pocock et al. [5] and the evaluation results are presented in Table 4.

Table 4: Suitability criteria.

Criteria	Suitability
Clarity of the aim	
Importance of engagement	
Resources available	
Scale of sampling	
Complexity of the protocol	
Motivations of participants	

The results for resources and participant motivation warrant some explanation:

- Resources available – since an online database is used, there are certain costs associated with storage and use. Currently the development of the application is done within the realm of an academic license of a university. The development costs associated with the ESRI products used (App Studio® and ArcGIS® Online) are covered by an academic license. With the final release of the product a custodian needs to be appointed which will assume the responsibility of system, associated maintenance and license costs.
- Motivations of participants – the challenge with a borehole is that special equipment is required for measurements. The public, however, have the ability to obtain the GPS location, take a photo and answer some basic questions in the form of selection lists. The project will resonate with the general public for typical reasons like a sense of jeopardy (“my water resources are under threat”) or being part of a narrative (“I’m taking part with others”). In contrast to the public participants, professional participants will gain the most value from the project and motivation is a non-issue.

The project potential as citizen science project was also evaluated by applying the decision framework suggested by Pocock et al. [5]. The evaluation of the project potential was done considering only the general public, with no measuring equipment. The result indicated that the project has very good potential for a mass participation citizen science approach.



6.2 Implementation of mobile app

The high-level implementation design of the mobile app is depicted in Fig. 8 and two user types are distinguished in the design namely “general public” and “professionals”. The major distinction between these two user types is the ability to perform proper borehole measurements. All available data capture fields are visible to users. Users registered as professionals will be able to capture all fields as opposed to the general public users who is only presented with the ability to enter data through selection lists to determine the status of the borehole. The latter is implemented, not to overwhelm the general public users to keep their interest in locating boreholes.

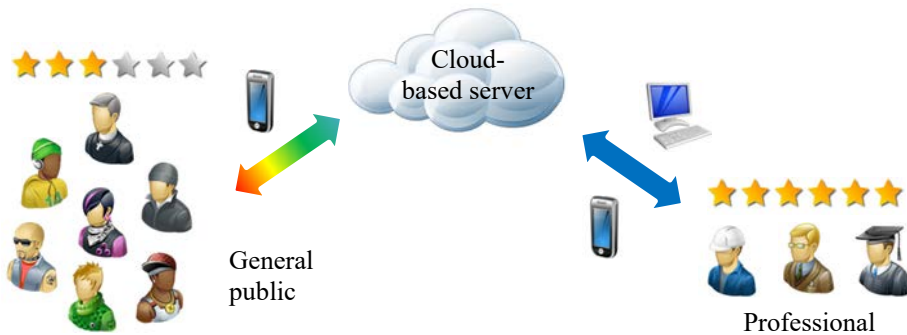


Figure 8: High level implementation of mobile app.

Reliable data which relates to an improved database are two of the identified success factors. The app implements the functionality where users verify each other's data. Each user is issued a user rating with registration and is issued a one-star rating. Data captured by a one-star user implies the data is low confidence until verified by a user of a higher rating. Multiple verifications lead to an increase of the star rating of a particular user.

One of the challenges of the existing NGA database is that users query the database, but seldom upload newly acquired data, due to the fact that the data has monetary value and they see this a competitive advantage in the groundwater industry when it comes to tenders.

The premise of the mobile app system is that only users who contribute to the database will have access to bulk downloads of data. This process is managed by means of a virtual credit system which is transparent to the users. Each user receives a number of credits during data update transactions and these credits are exchanged for download transactions. The credits have no monetary value, but rather transactional value.

A summary of how the success factors specified by Garcia-Soto et al. [4] were implemented or considered is presented in Table 5.

6.3 Pilot study area

The mobile app was tested on a pilot study area (Potchefstroom, South Africa), by making it available to people interested in participating. The user group consisted of ordinary citizens living in Potchefstroom, university students and a local environmental consultant.

At the start of the pilot, the available data from the online database (mainly NGA data as GRIP does not apply to the study area) amounted to 46 boreholes. The spatial distribution of

Table 5: Summary of success factors considered.

Success factor	Implementation/consideration
Clear goal	Determine borehole localities regardless if other borehole parameters can be measured by specific user.
Engagement of citizens	Engagement of geohydrologists is a non-issue, however, to keep the general public interested, outside the realm of a water crises, rely on factors like being part of a narrative: “I’m taking part with others”.
Reliable data	Data verification is achieved by users, for users through a rating scheme.
Improved database	By capturing actual borehole positions and associated data will result in an improved database.
Contribute to science	An improved database with both spatial and temporal data will support future data mining for trend analysis to be used in management.
Good communication	Training of users is funded by the Water Research Commission of South Africa – Project K5/2827.

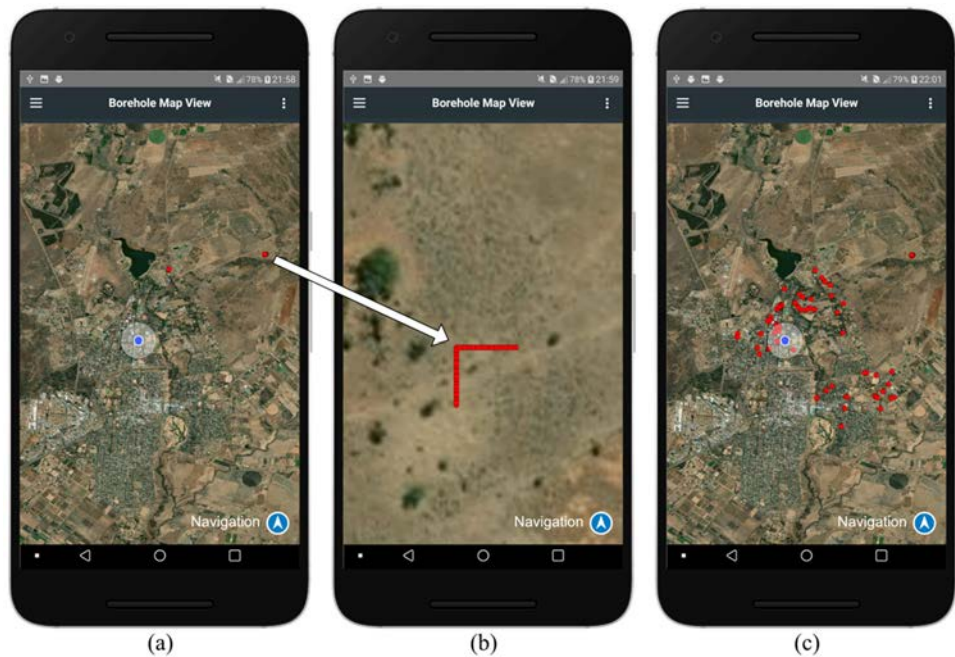


Figure 9: Map view of pilot study area. (a) Boreholes before pilot study; (b) Boreholes assigned to same location; and (c) Status at time of publication.

these 46 boreholes are shown in Fig. 9(a) and 9(b) respectively, and it is clear that the majority of the boreholes were historically assigned to the farm centroid at the time.

At the time of the publication and additional 63 borehole positions were recorded which totals 109 boreholes in the area as shown in Fig. 9(c). The borehole additions from the

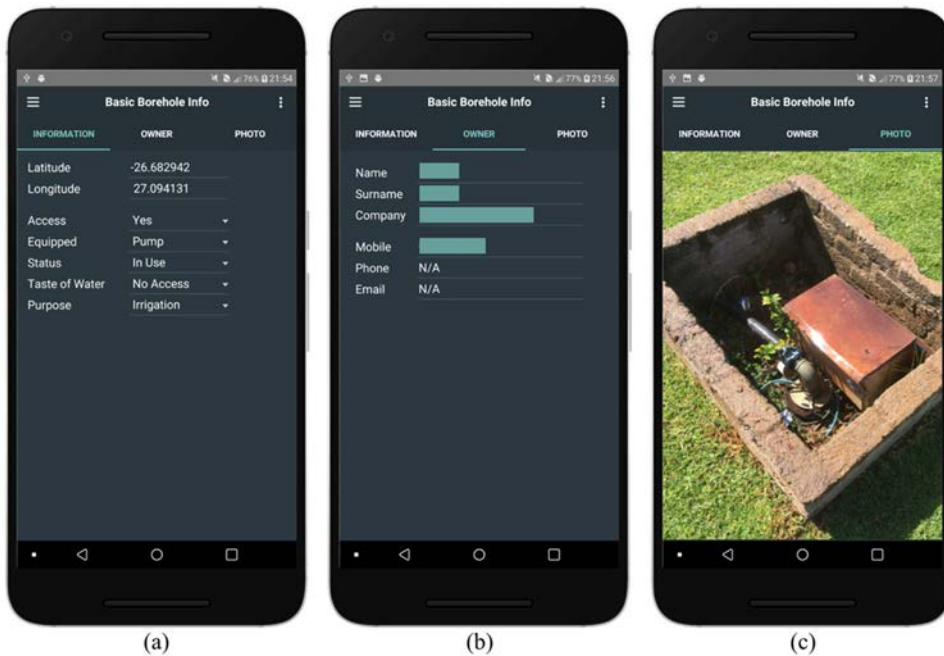


Figure 10: Basic borehole information. (a) Locality and status; (b) Owner; and (c) Photo.

different users were 33 from the general public, 19 from the student group and 11 from the local consultant. None of the existing boreholes (before commencement of the pilot study) were verified by any of the participating users. The consultant provided water level readings where the boreholes were not equipped. The general public users together with the student group only provided basic information as shown in the example shown in Fig. 10.

As mentioned earlier, even if only the borehole position can be verified, it is considered valuable information since locating a borehole in the field for hydro-census purposes is time consuming. Once the localities are known, professional users can carry out proper measurements to be used for management purposes and all users contributing to the database have access to this information.

7 CONCLUSIONS AND RECOMMENDATIONS

Both the assessment of the suitability and project potential were considered favourable, before the commencement of the actual software design. By choosing the App Studio® mobile application development environment from ESRI, a geo-enabled cross-platform native application could be built from a single code base.

The results from testing the app in a pilot study area, showed that borehole positions nearly doubled from what was on the existing database and 52% of these added borehole positions came from the general public. When considering the exiting boreholes before the pilot study, 96% of the boreholes were assigned to the same locality, making it impossible to locate them in the field. By having the public identify boreholes not in the database, could result in consolidating the “new” position with that of a “misplaced” borehole if enough information can be gathered to support this association. This however was not the case with the pilot

study. The pilot study is considered a success when considering the addition of new borehole localities that is verifiable in the field.

It is recommended to also include rainfall monitoring in the app as farmers in South Africa tend to keep meticulous rainfall records as their livelihoods depend on this.

ACKNOWLEDGEMENT

The Water Research Commission of South Africa is acknowledged for funding the project entitled “Mobile App for Hydrocensus and Groundwater Monitoring” (WRC Project K5/2827).

REFERENCES

- [1] SciStarter, Science we can do together. www.scistarter.org/citizen-science. Accessed on: 25 Aug. 2019.
- [2] Irwin, A., No PhDs needed: How citizen science is transforming research. *Nature*, **562**, pp. 480–482, 2018.
- [3] Stoop, J., New ways to use smartphones for science. www.elsevier.com/connect/new-ways-to-use-smartphones-for-science. Accessed on: 23 Aug. 2019.
- [4] Garcia-Soto, C. et al., Advancing citizen science for coastal and ocean research. Position Paper 23 of the European Marine Board, Ostend, Belgium, 2017.
- [5] Pocock, M.J.O., Chapman, D.S., Sheppard, L.J. & Roy, H.E., Choosing and using citizen science: A guide to when and how to use citizen science to monitor biodiversity and the environment. Centre for Ecology and Hydrology, 2014.
- [6] Ziervogel, G., Unpacking the Cape Town drought: Lessons learned. Report for Cities Support Programme, African Centre for Cities, 2019.
- [7] City of Cape Town, Day Zero: When is it, what is it and how can we avoid it? www.capetown.gov.za/media-and-news. Accessed on: 28 Aug. 2019.
- [8] Cape Town’s Water Crisis, Monitoring information and data about South Africa’s worst drought in history. www.capetowndrought.com. Accessed on: 28 Aug. 2019.
- [9] WRC, Geo-statistical analysis and sub-delineation of all Vegter Regions. Water Research Commission project number K5/2745, 2019.

