Citizen Science data integration SDG 6.3.2

WWQA World Water Quality Alliance

REPORT: Opportunities and challenges of integrating citizen sciencegenerated data on water quality for monitoring and achieving Sustainable Development Goal 6.3.2

(Specifics from Kenya & South Africa)

Prepared by Claire Williams[1], Steven Loiselle[1], Nicholas B. Pattinson[2]

[1]Earthwatch Europe, UK

[2] GroundTruth, Hilton, KwaZulu-Natal, South Africa

KENYA

The Kenya Water Resource User Associations (WRUAs) are key stakeholders that link the national Water Resource Agency to the water bodies and communities that depend on them. For more than 10 years, several WRUAs have been conducting water quality monitoring using kits funded through a German GIZ grant. These 40 WRUAs began monitoring important rivers and lakes in various areas of Kenya in 2012. Unfortunately, the data were stored on individual computers, resulting in data loss. Within this project, we developed an online database and data acquisition app for the WRUAs already engaged in monitoring, and we expanded monitoring to 20 additional WRUAs. Data from 2019 has been transferred to this database, which will be linked to the monitoring initiated using the FWW approach.

The key challenge that remains is to expand monitoring from these WRUAs to a larger number of the 400 WRUAs in Kenya, initially focusing on those most critical

SOUTH AFRICA

South Africa faces severe freshwater quality issues, tied primarily to poor wastewater management and large areas of agriculture. National government monitoring is infrequent and at a low temporal resolution, creating insufficient data to inform effective water resource management strategies. In response, a research project, termed the citizen science State of Rivers, is underway to use citizen science for monitoring and reporting on water quality in rivers across the country. Within the World Water Quality Alliance (WWQA) citizen science data integration for Sustainable Development Goal (SDG) 6.3.2 reporting project, >30 individuals and organisations from the citizen science State of Rivers network in South Africa were sent Freshwater Watch water quality monitoring kits and trained in their use. The network then gathered data between July and September 2024, ultimately collecting >120 comprehensive Freshwater Watch surveys on a variety of waterbodies across the country. These Freshwater Watch data were uploaded to the Freshwater Watch platform, contributing directly to a global findable, accessible, interoperable, and reusable (FAIR) database useful in SDG 6.3.2 reporting. However, data collected by other means, including using the mini stream assessment scoring system (miniSASS) were captured using other mechanisms, leading to some data not entering the SDG 632 reporting stream. There is a strong citizen science community of practice in South Africa focusing on water quality issues, creating excellent opportunity for citizen science data to augment national and international monitoring schemes.

OPPORTUNITIES

Combining citizen science data from multiple sources offers several exciting opportunities:

- **Enhanced Data Accuracy and Reliability**: Aggregating data from diverse sources can help cross-verify information and reduce the impact of individual errors or biases. It can provide a more comprehensive and accurate picture of the phenomena being studied.

  In South Africa, citizen scientists collect data using multiple approaches at the same site (for example, Freshwater Watch kits, miniSASS, clarity tubes, and Hach© nitrate test strips), often while professional scientists use more comprehensive techniques (for example, the South African Scoring System (SASS) version 5 or laboratory tests) allowing for different methods to be compared and validated against one another and against rigorous scientific methods.

  In Kenya, there are a number of projects that years of activity, with funding from multiple sources. In several cases, the measurements are being made using the standard FWW methodology, with the related quality control and database. The WRUA monitoring uses calibrated sondes for dissolved oxygen, conductivity and pH. The creation of the WRUA database anhd app will allow for better data management and timnely quality control. Several projects plan to utilise the miniSASS approach in the coming months.

- **Broader Geographic Coverage**: Citizen science projects often have participants from various locations. Combining data from multiple projects can cover a wider geographic area than possible via conventional means.

  In South Africa, the mandate of water quality monitoring falls to the governmental Department of Water and Sanitation (DWS), which has monitoring sites across the country. These sites are sparse and infrequently monitored, leaving large reaches of key rivers and other waterbodies unmonitored. The citizen science State of Rivers monitoring joint with this project brought together a large community of practice comprising >30 individuals and organisations, allowing for better coverage at a high

CHALLENGES

Combining citizen science data from multiple sources comes with challenges:

- **Data Quality and Consistency**: Citizen science data can vary widely in quality. Different contributors might use different methods, equipment, or have different levels of expertise.
  Ensuring that all data meets a consistent standard can be difficult.

  Because citizen science data are collected by multiple groups in different places, often using different methods or with different levels of training, ensuring that data are comparable and useable in an integrated database is a key challenge. Standardised training of a particular rigor and regular engagement to follow-up on approach consistency are required to ensure data accuracy, though the logistics of achieving such engagement can be difficult.

- **Data Validation and Verification**: A core component of data quality is that the data are verified and validated, which can be resource-intensive to achieve. A major barrier to data integration is ensuring that data from all sources are verified for accuracy and validated in their methodology.

- **Standardisation**: Aligning data collection methodologies to target similar phenomena can be vital to interrogating trends in metrics or events of interest, given that data collected using different criteria or classifications cannot form an interoperable, comparable singular database. Moreover, data from different sources might be formatted differently or use different units of measurement. Standardising these data (e.g., units for values, similar data presentation and storage styles, same data formats) so they can be effectively combined requires careful preprocessing and potentially significant effort.

- **Integration of Diverse Data Types**: Linked to standardisation, there can be significant challenges associated with the fact that citizen science data from different sources and groups can include text, images, numerical measurements, or other types of data. Integrating these diverse types into a coherent dataset can be complex.

- **Metadata and Documentation**: Comprehensive metadata (data about the data) are crucial for understanding the context in which data was

frequency, especially in areas where citizens experienced water quality issues. In Kenya, the Water Resources Agency is tasked with water quality monitoring and reporting. However, the WRUAs are direct stakeholders and well placed for long term monitoring and using these data for resource management to improve water quality. The project has reinforced and expanded monitoring through the WRUAs as well as increased monitoring in the Nairobi River basin. • **Increased Sample Size**: Larger datasets increase statistical power and the ability to detect subtle patterns or trends that might not be apparent in smaller datasets. During the project, >120 comprehensive Freshwater Watch surveys were completed and >70 miniSASS surveys were completed in South Africa, among other citizen science data. This shows the potential to greatly augment conventional monitoring databases and track water quality more closely, especially in impacted areas where concerned citizens can monitor frequently. Within the multiple projects in Kenya, more than 50 datasets have been obtained (FWW) or secured (WRUA data). The integration of this data presents a major improvement in the spatial and temporal coverage of water quality information. • **Enhanced Temporal Resolution**: Citizen science data can provide data frequently and in response to events such as floods or pollution spills as they happen. This enables more continuous and detailed time series data, which are crucial for studying temporal changes and understanding longterm trends. In combination with improved geographic coverage, enhanced increased sample sizes, • **Cross-Disciplinary Insights**: Integrating data from different types of citizen science projects can reveal new insights that are not apparent within a single discipline. Bringing together the community of practice in South Africa allowed diverse people in different locations to become aware of each other and coordinate their efforts using a unified technique. This allowed comparable	collected. Different sources might have varying levels of metadata or documentation, which can complicate integration and interpretation. Insufficient metadata to accompany citizen science data collected from different sources is a critical barrier to creating an integrated database. Without metadata (e.g., GPS coordinates, collection methodology, collector expertise level, site conditions), data integration is not possible. • **Spatial and Temporal Variability**: Data collected in different locations or times might need to be adjusted for spatial and temporal variations. This can be particularly challenging if data collection frequencies and times are inconsistent. The data from South Africa were collected over a good range of space and time. However, there was an evident bias for locations convenient to the citizens (e.g., within cities and near residential areas) or locations of key interest to citizens (e.g., reaches of rivers with known pollution issues). This geographic bias needs to be accounted for when considering the contribution of these data to an integrated database. The sampling was spread well through time, which was enabled by the fact that the South African community of monitors included many people part of official monitoring networks who sample each day of the week within ongoing monitoring regimes. Upscaling to citizens who are not able to contribute at certain times (e.g., weekends or weekdays) would be a key challenge to address. In Kenya, there are still major spatial gaps that need to be filled to allow for an understanding of the conditions of the countries water resources. This can only be met through a more extensive and long term commitment to citizen scientist activities. • **Data Privacy and Ethics**: Combining data from various sources raises concerns about privacy and consent, especially if the data includes sensitive information about individuals or locations. This was not an issue for data collected and integrated via the Freshwater Watch platform but presents a significant challenge to integrating more datasets: people contributing data need to know how their potentially private information might be securely stored and how their data might be used.

data to be collected by previous disparate groups, but also knowledge sharing on different approaches to collectively develop a shared, co-developed best practice for monitoring. This also provided a more comprehensive picture of water quality issues faced by different communities across the country than would have been possible from disconnected monitoring efforts and non-integrated data.

Bringing together different systems (FWW and WRUA), in Kenya has allowed for new insights into local pollution hotspots, areas of elevated water quality and seasonal trends, fundamental to identify possible pollution sources.

- **Greater Public Engagement**: Collaborative projects that merge citizen science data from multiple sources can engage a broader audience, encouraging more people to participate and contribute. This can also help build a stronger community around scientific research. This project brought together citizen science individuals, initiatives, and organisations, creating a community of practice all aware of one another and fostering improved engagement between groups.

- **Resource Efficiency**: Combining data can optimise the use of resources by minimising duplication of effort. This is apparent in this project through the contribution of different initiatives, teams, and individuals to a centralised database requiring centralised, coordinated curation for all data. This minimises redundant data handling when attempting to analyse data from disparate databases. This project showed how the data collection of different groups could be centrally coordinated through the Freshwater Watch mobile app and database. This project also illustrated how data integration can increase the visibility and power of existing datasets. By creating integrated data, it is possible to perform new analyses and generate new knowledge without requiring additional data collection.

- **Policy and Decision-Making Support**: Data integrated into a single, credible and recognised dataset can provide valuable evidence for policymakers and decision-makers, leading to more informed and effective policy decisions.

- **Coordination and Collaboration**: Coordinating between different citizen science projects or platforms can be challenging, especially if there are no established protocols for data sharing and integration.

  In South Africa, data are collected by numerous groups using different platforms. This results in data of varying formats and contents being captured in various platforms and datasets. Great effort is required to coordinate among groups to gather data into a collective database for analyses or comparison. Typically, this needs to happen through manual effort of someone coordinating which might not be available or feasible.

  The present project has demonstrated the opportunities to bring together two major projects and datasets. However, the transcription of past data from paper logs, or from excel sheets to an open and secured database requires time. The required funding for such transcription and validate of past data is not always available.

- **Technical and Computational Issues**: Merging large datasets from various sources can involve significant computational resources and technical expertise, especially when dealing with big data. This project highlighted that the local infrastructure for data capture can be a considerable issue. Data integration happens via smart technology and reliable internet connections. In developing regions, especially in rural settings, access to technological devices and internet can be a real challenge for data capture and submission.

- **Financial barriers**: Though citizen science is typically far more cost-effective than conventional monitoring, it can still require significant financial investment from participants. In this project, the need for financial aid in procuring and supplying data collection kits to participants at no cost to them was highlighted as a challenge for scaling. Moreover, the project emphasised that financial barriers can be a significant barrier to participation. This needs to be considered when aiming for inclusivity goals, where many people will be financially excluded from partaking in citizen science, even at low cost and especially in developing regions.

This was evident in this project through the contribution of all the data collected via the Freshwater Watch app going into the Freshwater Watch database, a key resource for informing policy documentation and recommendations to governments and managers.	

Relevant references:

August, T.A., West, S.E., Robson, H., Lyon, J., Huddart, J., Velasquez, L.F. and Thornhill, I., 2019. Citizen meets social science: predicting volunteer involvement in a global freshwater monitoring experiment. *Freshwater science*, *38*(2), pp.321-331. https://doi.org/10.1086/703416

Bishop, I.J., Warner, S., van Noordwijk, T.C., Nyoni, F.C. and Loiselle, S., 2020. Citizen Science Monitoring for Sustainable Development Goal Indicator 6.3. 2 in England and Zambia. Sustainability, 12(24), p.10271. https://doi.org/10.3390/su122410271

Di Grazia, F.; Gumiero, B.; Galgani, L.; Troiani, E.; Ferri, M.; Loiselle, S.A. 2021. Ecosystem Services Evaluation of Nature-Based Solutions with the Help of Citizen Scientists. Sustainability, 13, 10629. https://doi.org/10.3390/su131910629

Graham, P.M., and Taylor, J. 2018. Development of citizen science water resource monitoring tools and communities of practice for South Africa, Africa and the world. P.M. Graham & J. Taylor (Eds.). Water Research Commission (WRC) Report No. TT 763/18, Pretoria, South Africa. Available at https://www.wrc.org.za/wpcontent/uploads/mdocs/TT%20763%20web.pdf

Graham, P.M., Pattinson, N.B., Lepheana, A.T., and Taylor, J. (2024). Clarity tubes as effective citizen science tools for monitoring wastewater treatment works and rivers. Integrated Environmental Assessment and Management, 20(5), 1463-1472. https://doi.org/10.1002/ieam.4937

Hadj-Hammou, J., Loiselle, S., Ophof, D., Thornhill, I., 2017. Getting the full picture: Assessing the complementarity of citizen science and agency monitoring data, PloS one 12 (12), e0188507 https://doi.org/10.1371/journal.pone.0188507

Hegarty, S., Hayes, A., Regan, F., Bishop, I. and Clinton, R., 2021. Using citizen science to understand river water quality while filling data gaps to meet United Nations Sustainable Development Goal 6 objectives. *Science of The Total Environment*, *783*, p.146953

Lepheana, A.T., Russell, C., and Taylor, J. 2021. Co-researching transformation within training processes in a post COVID-19 world: The case story of the Palmiet Enviro-Champs, indigenous knowledge practices and Action Learning. In I. Kulundu-Bolus, G. Chakona, & H. Lotz-Sisitka (Eds.), Stories of collective learning and care during a pandemic: Reflective research by practitioners, researchers and community-based organisers on the collective shifts and praxis needed to regenerate transformative futures (pp. 55–82). Transforming Education for Sustainable Futures (TESF) and the Rhodes University (RU) Environmental Learning Research Centre (ELRC). https://doi.org/https://doi.org/10.5281/zenodo.5704833

Loiselle, S., Bishop, I., Moorhouse, H., Pilat, C., Koelman, E., Nelson, R., Clymans, W., Pratt, J. and Lewis, V., 2024. Citizen scientists filling knowledge gaps of phosphate pollution dynamics in rural areas. Environmental Monitoring and Assessment, 196(2), pp.1-15. https://doi.org/10.1007/s10661-024-12389-5

Lotz-Sisitka, H., Ward, M., Taylor, J., Vallabh, P., Madiba, M., Graham, P.M., Louw, A.J., and Brownell, F. 2022. Alignment, scaling and resourcing of citizen-based water quality monitoring Initiatives. Water Research Commission (WRC) Report No. 2854/1/22, Pretoria, South Africa. Available at https://www.wrc.org.za/wpcontent/uploads/mdocs/2854%20final.pdf

Miguel-Chinchilla, L., Heasley, E., Loiselle, S., Thornhill, I. 2019. Local and landscape influences on turbidity in urban streams: a global approach using citizen scientists Freshwater Science, 38(1), https://doi.org/10.1086/703460

Moshi, H.A., Kimirei, I., Shilla, D., O'Reilly, C., Wehrli, B., Ehrenfels, B. and Loiselle, S., 2022. Citizen scientist monitoring accurately reveals nutrient pollution dynamics in Lake Tanganyika coastal waters. Environmental monitoring and assessment, 194(10), pp.1-18. https://doi.org/10.1007/s10661-022-10354-8

Moshi, H.A., Shilla, D.A., Brehim, J., Kimirei, I., O'Reilly, C. and Loiselle, S., 2023. Sustainable Management of the African Great Lake Coastal Areas: Motivations and Perspectives of Community Citizen Scientists. Environmental Management, pp.1-15. https://doi.org/10.1007/s00267-023-01824-x

Moshi, H.A., Shilla, D.A., Kimirei, I.A., O'Reilly, C., Clymans, W., Bishop, I. and Loiselle, S.A., 2022. Community monitoring of coliform pollution in Lake Tanganyika.
PloS one, 17(1), p.e0262881. https://doi.org/10.1371/journal.pone.0262881

Pattinson, N.B., Dickens, C.W.S., Taylor, J., and Graham, P.M. 2024. Smartphones for citizen science water quality monitoring in developing regions. Colombo, Sri Lanka: International Water Management Institute (IWMI). CGIAR Initiative on Digital Innovation. 22p. Available at https://hdl.handle.net/10568/151938

Pattinson, N.B., Maharaj, U., Singh, K., Taylor, J., Lepheana, A.T., Dickens, C.W.S., and Graham, P.M. 2024. Digitally enhanced community-based environmental monitoring: Technologically upgrading the Enviro-Champs initiative. Colombo, Sri Lanka: International Water Management Institute (IWMI). CGIAR Initiative on Digital Innovation. 16p. Available at https://hdl.handle.net/10568/151937

Pattinson, N.B., Russell, C., Taylor, J., Dickens, C.W.S., Koen, R.C.J., Koen, F.J., and Graham, P.M. 2023. Digital innovation with miniSASS, a citizen science biomonitoring tool. Colombo, Sri Lanka: International Water Management Institute (IWMI). CGIAR Initiative on Digital Innovation. 31p. https://hdl.handle.net/10568/134498

Pattinson, N.B., Taylor, J., Dickens, C.W.S., and Graham, P.M. 2023. Digital innovation in citizen science to enhance water quality monitoring in developing countries. Colombo, Sri Lanka: International Water Management Institute (IWMI). CGIAR Initiative on Digital Innovation. 40p. (IWMI Working Paper 210). https://doi.org/10.5337/2024.201

Pattinson, N.B., Taylor, J., Lepheana, A.T., Dickens, C.W.S., and Graham, P.M. 2023. The enviro-champs: Establishing a framework for a technologically upgraded environmental monitoring network at community scale. Colombo, Sri Lanka: International Water Management Institute (IWMI). CGIAR Initiative on Digital Innovation. 19p. Available at https://hdl.handle.net/10568/138440

Taylor, J., Graham, P.M., Louw, A.J., Lepheana, A.T., Madikizela, B., Dickens, C.W.S., Chapman, D.V., and Warner, S. 2022. Social change innovations, citizen science, miniSASS and the SDGs. Water Policy, 24(5), 708–717. https://doi.org/10.2166/wp.2021.264

19. Thornhill, I., Loiselle, S., Clymans, W. and van Noordwijk, C.G.E., 2019. How citizen scientists can enrich freshwater science as contributors, collaborators, and co-creators. Freshwater Science, 38(2), pp.231-235. https://doi.org/10.1086/703378

Thornhill, I.; Chautard, A.; Loiselle, S. 2018. Monitoring Biological and Chemical Trends in Temperate Still Waters Using Citizen Science. Water, 10, 839. https://doi.org/10.3390/w10070839

Warner, S., Blanco Ramírez, S., de Vries, S., Marangu, N., Ateba Bessa, H., Toranzo, C., Imaralieva, M., Abrate, T., Kiminta, E., Castro, J. and de Souza, M.L., 2024.
Empowering citizen scientists to improve water quality: from monitoring to action. Frontiers in Water, 6, p.1367198. https://doi.org/10.3389/frwa.2024.1367198