USAID
FROM THE AMERICAN PEOPLE

GLOBAL WATER STRATEGY
ACTION RESEARCH INITIATIVE

LESSONS AND GOOD PRACTICES FOR BENCHMARKING FECAL SLUDGE MANAGEMENT

Final Report

June 2024

This report is made possible by the support of the American People through the United States Agency for International Development (USAID). The content of this report is the responsibility of The Aquaya Institute and Tetra Tech and does not necessarily reflect the views of USAID or the United States Government.

ACKNOWLEDGMENTS

The Aquaya Institute prepared this report. Haleemah Qureshi oversaw the Activity. Tom Mahin and Caroline Delaire carried out the literature review. Karen Setty led the case study research, key performance indicator (KPI) analysis, and final report compilation, relying on the generous knowledge contributions of key informants affiliated with the Faecal Sludge Management (FSM) Alliance, Sustainable Sanitation Alliance (SuSanA), German Development Agency (GIZ), Asian Institute of Technology (AIT) Thailand, Athena Infonomics, Eawag, Uganda's Ministry of Water and Environment, National Water and Sewerage Corporation (NWSC), Kampala Capital City Authority (KCCA), Sheercare Services, Makerere University, Zambia's National Water Supply and Sanitation Council (NWASCO), Lusaka Water Supply and Sanitation Company (LWSC), Tiyende Sanicon Limited, Water & Sanitation for the Urban Poor (WSUP) Zambia, WaterAid Bangladesh, SNV Bangladesh, CEPT University, the Indian Institute for Human Settlements, the Manila Metropolitan Waterworks and Sewerage System Regulatory Office, Islamic Development Bank (isDB), Fresh Life Kenya, Pit Vidura, and Sustainable Organic Integrated Livelihoods (SOIL) Haiti. On behalf of the United States Agency for International Development (USAID), Jesse Shapiro and Ryan Mahoney served as project advisors. On behalf of Tetra Tech and the Urban Resilience by Building and Applying New Evidence in Water, Sanitation, and Hygiene (URBAN WASH) consortium, Liz Jordan, Miriam Otoo, and Tanvi Nagpal served as technical advisors.

Preferred citation: USAID Urban Resilience by Building and Applying New Evidence in Water, Sanitation, and Hygiene (URBAN WASH). 2024. *Lessons and Good Practices for Benchmarking Fecal Sludge Management.* Washington, DC: USAID URBAN WASH project.

Prepared for the USAID by the URBAN WASH project, under the General Service Administration's (GSA's) One Acquisition Solicitation for Integrated Services (OASIS Unrestricted) Indefinite Delivery Indefinite Quantity Contract, contract number GS00Q14OADU138 and order number 7200AA21M00012.

Tetra Tech Contacts:

Liz Jordan, Chief of Party
liz.jordan@tetratech.com

Miriam Otoo, Deputy Chief of Party
miriam.otoo@tetratech.com

Zach Borrenpohl, Project Manager
zach.borrenpohl@tetratech.com

Tetra Tech
1320 N. Courthouse Road, Suite 600, Arlington, VA 22201
Tel: (703) 387-2100, Fax: (703) 414-5593
www.tetratech.com/intdev

Cover photo: Fecal sludge emptying truck in San Pedro, Cote d'Ivoire; photo courtesy of Tetra Tech

TABLE OF CONTENTS

ACRONYMS

ADB	Asian Development Bank
AIT	Asian Institute of Technology
AWWA	American Water Works Association
CWAS	Center for Water and Sanitation
CWIS	Citywide Inclusive Sanitation
DINEPA	National Directorate for Water and Sanitation
EBC	European Benchmarking Co-operation
ESAWAS	Eastern and Southern Africa Water and Sanitation Regulators Association
EWURA	Energy and Water Utilities Regulatory Authority
FSM	Fecal (or Faecal) Sludge Management
GIS	Geographic Information System
GIZ	German Development Agency
GLAAS	Global Analysis and Assessment of Sanitation and Drinking-Water
IBNET	International Benchmarking Network for Water and Sanitation Utilities
isDB	Islamic Development Bank
ITN-BUET	International Training Network Centre
IWA	International Water Association
JICA	Japan International Cooperation Agency
JMP	World Health Organization/United Nations Children's Fund Joint Monitoring Programme
KCCA	Kampala Capital City Authority
KPI	Key Performance Indicator
LWSC	Lusaka Water Supply and Sanitation Company
M	Million
NGO	Nongovernmental Organization
NWASCO	National Water Supply and Sanitation Council
NWSC	National Water and Sewerage Corporation
PAS	Performance Assessment System (India)
PWWA	Pacific Water and Wastewater Association

RURA	Rwanda Utilities Regulatory Authority
SOIL	Sustainable Organic Integrated Livelihoods Haiti
SuSanA	Sustainable Sanitation Alliance
UN	United Nations
UNICEF	United Nations Children's Fund
URBAN WASH	Urban Resilience by Building and Applying New Evidence in Water, Sanitation, and Hygiene
USAID	United States Agency for International Development
WASH	Water, Sanitation, and Hygiene
WASH-FIN	Water, Sanitation, and Hygiene Finance
WASREB	Water Services Regulatory Board
WHO	World Health Organization
WSP	Water and Sanitation Program
WSUP	Water & Sanitation for the Urban Poor

EXECUTIVE SUMMARY

Benchmarking involves assessing performance indicators that can be compared to targets, historical values, and other entities (e.g., persons, organizations, cities, or countries). This approach to fostering performance improvements contributes to the enabling environment for effective water and wastewater services, which have traditionally existed as non-competitive monopolies. Although large water and wastewater utilities have widely adopted benchmarking, the practice has only begun expanding to fecal sludge management (FSM), also called decentralized, on-site, or non-sewered sanitation. This report addresses the question: ***What are good practices for FSM benchmarking systems, and how should these be implemented in different institutional or governance contexts?*** The report synthesizes literature and case studies to identify enabling factors relevant to effective FSM benchmarking.

WHAT METRICS SHOULD FSM BENCHMARKING SYSTEMS MEASURE?

Multiple scales and types of benchmarking exist for water and sanitation providers, including organizational, citywide, national, regional, and international initiatives. Within-country benchmarking often relates to regulatory oversight, performance-based contracts, donor requirements, or voluntary initiatives. International benchmarking efforts, such as the World Bank-administered International Benchmarking Network for Water and Sanitation Utilities (IBNET) and the World Health Organization (WHO)/United Nations Children's Fund (UNICEF) Joint Monitoring Programme (JMP), primarily serve to guide development investments and to allow peer-to-peer comparisons.

"Performance" or "metric" benchmarking relies on monitoring quantitative key performance indicators (KPIs), most of which traditionally relate to commercial functions such as operations, productivity, and finances. "Process" or "practice" benchmarking supports peer learning and sharing good management practices, while other types include customer surveys and model company engineering. Recognizing the risk of neglecting social responsibility outcomes in performance data, newer benchmarking initiatives such as the Citywide Inclusive Sanitation (CWIS) Initiative have introduced KPIs on gender equity and social inclusion.

The diversity of actors involved in carrying out on-site sanitation services often leads to a multiplicity of KPIs, which makes data collection resource-intensive and interpretation by local officials difficult. Quantifying FSM KPIs requires collecting data from multiple entities, including households, desludging operators, treatment plant operators, and local government, as opposed to one utility for water and wastewater. This study recommends consistently applying a short list of high-priority indicators to compare city-level FSM. Service providers should then tailor additional indicators to local needs and aggregate a smaller numer of indicators for global comparison. The recommended indicators are:

- Percentage of population with access to safely managed sanitation,
- Access to sanitation services among vulnerable groups,
- Coverage of on-site sanitation facilities,
- Total collection volume,
- Treatment ratio (volume treated: collected),
- Percentage of fecal sludge safely managed,
- Worker safety procedure adherence,

- Operating cost recovery, and
- Percentage of desludging providers operating with a formal license.

WHAT FACTORS DRIVE ADOPTION AND USE OF BENCHMARKING FOR FSM?

A functioning FSM system must underpin the use of benchmarking to drive performance improvement, requiring attention to infrastructure, regulatory standards, tariffs, organizational readiness, and communication among participating actors. Building blocks that aid benchmarking startup include regulatory sensitization and requirements, a customer-engaged business model, financial support tied to benchmarking, and adequate data management capacity within and across stakeholder organizations.

Good practices expected to reinforce sustained effectiveness of water, wastewater, and FSM benchmarking programs revolve around selecting a limited number of consistent indicators, adequate data management systems, incentives, and public information sharing. Stakeholders can mix incentives, which may include direct or indirect financial rewards (e.g., performance-based funding, eligibility for grants), as well as reputational or regulatory penalties for poor performance or lack of participation in benchmarking programs. Public data sharing offers greater data accessibility when promoted by high-level buy-in, financial support, data validation and correction opportunities, confidentiality measures, and user-friendly data entry and access interfaces.

WHAT ARE THE BENEFITS OF BENCHMARKING FSM SERVICES?

Common-sense rationale drives widespread benchmarking application. Robust investigations to confirm that benchmarking improves service delivery performance can be challenging to conduct within real-world complexities, leading to few critical evaluations. Some criticize benchmarking as promoting narrow, top-down, and Western-centric views of performance that distract from local problem-solving and innovation. Still, the literature and case studies largely support directions taken by the relaunched "New IBNET" and others to reap the potential benefits of benchmarking activities while minimizing negative effects. Even given substantive contextual challenges, FSM benchmarking is a versatile tool that can improve transparency, foster competition, and inform advocacy and financial and management decisions.

1.0 INTRODUCTION

Benchmarking involves characterizing the performance of an organization by comparing indicator values to pre-defined standards or objectives or comparing among organizations or geographic areas. Comparative benchmarking allows service providers to learn from one another (Sharma 2006). It originated in the private manufacturing industry in the 1970s as a strategic tool to stay ahead of competitors (Murungi and Blokland 2016). The approach has since expanded to other public service industries, including the water and wastewater sector, allowing governments, regulators, funders, and others to compare performance across service providers, capture historical trends, and identify best practices (S. Berg and Padowski 2010). Clear comparators promote transparency and accountability, which are particularly important in a sector where service providers typically enjoy local monopolies, limiting opportunities for market competition to incentivize improvements (Sharma 2006; Lambert 2021). Thus, benchmarking represents one element enabling effective water and wastewater services (S. Berg 2010, 2020; Sinha 2013; URBAN WASH 2023).

Benchmarking activities can be formal or informal. Two primary types of benchmarking include "metric" benchmarking, which involves measuring performance through quantitative indicators, and "process" benchmarking, which involves improving practices through peer knowledge sharing and imitation, such as adopting specific protocols or guidance (IBNET 2021a; Blankenship, Olstein, and Liner 1998). These two types of benchmarking are complementary as metric benchmarking can lead to process benchmarking and promote directed learning and actions to achieve better outcomes. A number of other types of benchmarking exist, including customer surveys and model company engineering (Mann et al. 2010; IBNET 2021a).

This study focused on metric, also called performance, benchmarking, which entails comparing quantitative indicators to targets, historical values, and other service providers (AWWA 2020; S. Berg 2020). Key performance indicators (KPIs) are metrics that index operation, production, capacity, and cost outcomes, enabling decision-makers to track trends and identify needed changes. KPI values contain valuable information signifying performance improvement or worsening and may be used in designing regulatory and managerial programs. A "benchmark" value represents the reference or desired target level of the KPI (AWWA 2020).

Although widespread among water and wastewater utilities, benchmarking has not yet been widely applied to fecal (or faecal) sludge management (FSM), also called on-site, decentralized, or non-sewered sanitation. To safely dispose of fecal sludge accumulated on-site, different service providers need to coordinate to carry out various steps in the FSM chain, including sludge containment, collection, transport, treatment, and disposal or reuse. Inadequate FSM can lead to surface and groundwater contamination, as well as the spread of disease. Demographic and Health Survey data from 58 countries over the period of 2003–2015 established that 63 percent of households use sanitation facilities requiring FSM, totaling approximately 1.8 billion people (Berendes, Sumner, and Brown 2017). In low-resource countries, fecal sludge is often unsafely managed, with a larger proportion discharged directly into the environment or waterways (Sustainable Sanitation Alliance 2023; WSUP 2019).

Applying benchmarking to FSM could reinforce progress toward public and environmental health goals. To inform future implementation research programs, this study sought to address the question: ***What are good practices for FSM benchmarking systems, and how should these be implemented in different institutional or governance contexts?*** Due to the limited activity and evidence on FSM benchmarking, the study synthesized learnings from:

1. Literature on benchmarking, largely from the water and wastewater sector; and
2. Case studies in locations that have developed FSM benchmarking frameworks and/or pioneered FSM benchmarking.

The following research sub-questions guided the methods and analysis for the literature review and case studies:

1. What are good practices around benchmarking from the water and wastewater sector, and to what extent can those translate to FSM or citywide inclusive sanitation (CWIS) benchmarking systems?
2. What data, monitoring, scoring, and incentive approaches are needed to support an effective, transparent, and sustainable benchmarking system?
3. What factors drive adoption of benchmarking for FSM and foster use of KPIs as a standard part of service provider operations?
4. What factors support (or limit) FSM benchmarking systems in different institutional or governance contexts?
5. What range of outcomes can be expected from implementing a benchmarking system for FSM?
6. At what scale (global, regional, national, subnational) and level of aggregation (e.g., outcomes for individual service providers versus cities as a whole) are benchmarking systems most effective at encouraging healthy competition and supporting performance improvements?
7. Which KPIs should stakeholders prioritize to improve FSM performance?

2.0 METHODS

2.1 LITERATURE REVIEW

In the first research phase, the Urban Resilience by Building and Applying New Evidence in Water, Sanitation, and Hygiene (URBAN WASH) team conducted a narrative literature review (Ferrari 2015). Because national-level FSM benchmarking applications have only emerged since 2014, limited documentation addresses FSM benchmarking specifically. Thus, we searched for relevant articles, reports, manuals, and policies using a variety of search terms inclusive of water, sewered sanitation, and non-sewered sanitation services (e.g., benchmark water, benchmark sewer, benchmark FSM, KPIs, KPI water, KPI sewer, KPI sludge) on Google, Google Scholar, and PubMed.

We also carried out limited searches of the literature on benchmarking for solid waste management, electricity, and healthcare to identify practices that may be relevant to FSM. We paid specific attention to understanding how benchmarking may differentially affect women, youth, or marginalized groups. This involved identifying examples of benchmarking promoting gender equity and pro-poor service delivery, as well as potential unintended consequences of benchmarking or biases disadvantaging vulnerable groups.

From these searches, we identified 54 peer-reviewed journal articles and more than 100 gray literature sources. The information reviewed included studies from Africa, Asia and the Pacific, Latin America and the Caribbean, Europe, and North America.

2.2 CASE STUDIES

In the second research phase, case study key informant interviews offered a means to collect up-to-date information from nascent FSM benchmarking efforts. We used the findings to augment the literature review and synthesize knowledge from both sources.

2.2.1 SELECTION AND RECRUITMENT

We nominated country case study options (n = 30; Appendix A) through:

- The initial literature review,
- A survey of urban FSM in Africa (JICA 2022),
- A review of FSM data use cases (The Aquaya Institute 2022b),
- A Tableau tool digitizing the Sustainable Sanitation Alliance (SuSanA) shit-flow diagram data (The Aquaya Institute 2022a; Sustainable Sanitation Alliance 2023),
- A review of CWIS pioneers featured by the World Bank (World Bank 2023), and
- Web scanning for FSM benchmarking activity using artificial intelligence (AI) technology (OpenAI 2021) followed by human verification.

When screening locations, we aimed to capture a spectrum of implementation progress ranging from pilot FSM benchmarking programs in one urban area to mandated national expansion, using the programming origination year, scale, and number of KPIs as proxies for the stage of progress. The search criteria favored:

- Low- or lower-middle-income countries,

- A large portion of the population using on-site sanitation and FSM services,
- FSM-tailored institutional and/or regulatory frameworks,
- Documented FSM benchmarking systems in place for at least one year, and/or
- A variety of actors driving progress in FSM: utility, municipality, or regulator.

We excluded short-listed case study locations if they lacked documentation of active FSM benchmarking activities (n = 20) or access to experts we could interview (n = 2; Appendix A). To recruit key informants, we initially reached out via email to existing contacts within the prospective case study locations, including former collaborators and URBAN WASH advisory board members. We used snowball sampling to request referrals to additional contacts, aiming for a diversity of perspectives. In addition, we created an open recruitment flyer and shared this with a professional FSM network on WhatsApp. Interviewees included government regulators, service providers (e.g., utilities, emptiers), NGOs, and educational institutions (Appendix A).

2.2.2 CASE STUDY COMPARISON

The FSM benchmarking case study sample, as intended, captured varied characteristics. Table 1 shows the status of FSM policies, benchmarking activity, influential actors, and the scale of data collection (approximated through the number of KPIs) for selected country case studies. Other information, such as country income and world region diversity, is captured in Appendix A, Table A1. The number of KPIs serves as a proxy for the scale and direction of movement of benchmarking activities, which the literature and key informants reflected as undergoing change across most locations.

Table 1: Case study characteristics related to benchmarking activity for FSM, including policy and activity status, key actors, and the number of KPIs in use

Country	National FSM Policy Status	FSM Benchmarking Activity Status	Key Actors	# KPIs[1]
Uganda	No policy	Implementation by Kampala city authority with donor support	Kampala Capital City Authority (KCCA), National Water and Sewerage Corporation (NWSC), Gates Foundation, German Development Agency (GIZ)	20–40↓
Zambia	National framework established	Tracking one indicator nationally and planning scale-up using Lusaka city example	Lusaka Water Supply and Sanitation Company (LWSC), National Water Supply and Sanitation Council (NWASCO), Gates Foundation, Water & Sanitation for the Urban Poor (WSUP)	1 → 3+
Bangladesh	National policy established	National dashboard in progress	Ministry of Local Government, International Training Network Centre (ITN-BUET), Asian Institute of Technology (AIT), international nongovernmental organization (iNGOs)	~8
India	National policy established	FSM indicators recommended nationally but not yet fully integrated with local sewerage	CEPT University, Center for Water and Sanitation (CWAS), Gates Foundation, Ministry of Urban Development	1 → 6 → 16?

Country	National FSM Policy Status	FSM Benchmarking Activity Status	Key Actors	# KPIs[1]
		reporting within select states		
Indonesia	No policy	Implementation at donor-supported project level	Islamic Development Bank (IsDB), Gates Foundation, United States Agency for International Development (USAID), Asian Infrastructure Investment Bank (AIIB), Borda	1–18
Philippines	National policy established, national regulator pending	Implementation at privatized service provider level with city-level regulatory oversight	Manila Water, Maynilad, Metropolitan Waterworks and Sewerage System (MWSS) Regulatory Office	~3
Kenya	National policy pending	Information system and data hubs under development with donor support	Ministry of Health, United Nations Children's Fund (UNICEF), USAID, Government of Japan	1↑
Haiti	No policy	Informal implementation at social enterprise and project level, operating with donor support	Sustainable Organic Integrated Livelihoods (SOIL) Haiti, National Directorate for Water and Sanitation (DINEPA)	10–40

[1]Arrows, ranges, and symbols represent ongoing transitions to increase or decrease the number of KPIs in use specifically for on-site sanitation.

2.2.3 DATA COLLECTION

The interview guide (Appendix B) covered FSM and FSM benchmarking approaches, actors, strengths, gaps, and future outlook. The interview questions and data analysis categories were developed using the Sanitation Social-Ecological System framework (Trimmer, Miller, et al. 2020; Trimmer, Lohman, et al. 2020), which considers the roles of resources, services, technologies, actors, and governance structures and their interactions within a broader social, economic, political, and environmental context. After learning during initial outreach that the "benchmarking" topic (and terminology) was not yet advanced or similar across contexts, we added an introductory question sharing our definition of benchmarking and asking how each respondent defined it.

Between July and August 2023, we reached out to approximately 70 individual contacts, resulting in a total of 22 virtual interviews with 27 individuals, a response rate of approximately 31 percent (Table 2). Seventeen responses relayed experiences from one country, while five addressed activities in more than one country. Males constituted two-thirds (17/27) of respondents. Key informants had the option to respond to the interview questions in writing or provide verbal responses recorded by an interviewer during a live video call. One organization submitted only written inputs due to a schedule conflict, one submitted written inputs prior to the live interview, and one provided additional written inputs following the live interview from a colleague who could not attend. In addition, respondents in many cases referred us to documents specific to the case study locations.

Given the availability of respondents, we conducted:

- Two full case studies for depth,
- Two partial case studies, and

- Four brief case studies for breadth (Table 2).

Table 2: Response for key informant interviews by location

Information Use	Country*	Interview Responses**
Full case studies for depth	Uganda	5
	Zambia	4
Partial case studies	India	2
	Bangladesh	2
Brief case studies for breadth	Indonesia	1
	Philippines	1
	Kenya	1
	Haiti	1
Global context	Global or multiple countries	5
Total		**22**

*Cases were selected at the national level, although the geographic range of FSM benchmarking activity may have been limited to one or more urban centers.

**Some interviews involved more than one person, and some respondents provided information on more than one country.

2.2.4 DATA ANALYSIS

We reviewed and extracted written interview notes into categories (Table 3) for analysis, formulating a data matrix by case study location and topic. For each case study, we documented the FSM context, benchmarking approaches, and specific issues, including KPIs used for FSM, opportunities for future growth, and benchmarking effects. Additional documentation supplied by respondents for specific case study locations (approximately 45 resources) supplemented the data matrix and written report.

Table 3: Categories applied to analyze interview data from FSM case studies

Background Context	FSM Benchmarking Approach	Specific Issues
• General FSM approach • Actors • Regulation and governance • Sociopolitical and economic context • Incentives	• Origins • Type of benchmarking • KPIs • Process or other active benchmarking • Change process • Effects of performance tracking • Future priorities	• Data quality • Data reporting • Data use • Data coverage of sanitation service chain • Gender and social inclusion • Climate and resources

While synthesizing the interview data, we drafted memos to record inductive observations alongside supporting evidence. We then grouped the memos to address specific research questions. Working across case studies, we:

- Observed parallels and contrasts;
- Distilled good practices surrounding data quality, data dissemination, and incentives; and
- Highlighted strategies that could augment FSM benchmarking or expand it into new settings for future sustainable development programming.

Using the case study interview data, we also developed a diagram summarizing the conditions, building blocks, process steps, and corresponding pitfalls affecting successful FSM benchmarking implementation,

(USAID Learning Lab 2022). We validated this summary of findings with a subset of respondents by sharing the diagram with one organization per country to confirm that it properly reflected their understanding of the case study conditions. We received four responses out of eight requests: two confirming the summary findings and two offering additional clarifications.

2.3 KEY PERFORMANCE INDICATOR PRIORITIZATION

To formulate recommendations regarding FSM-specific KPIs, we collated a list from multiple sources, including the literature review and case studies. We first grouped the indicators into relevant categories. For FSM, representation of indicator categories across the sanitation service chain captured practical information about different components involved in FSM services (Strande, Ronteltap, and Brdjanovic 2014). In addition, we added categories that captured other elements critical to the FSM business model and global development goals, which some argue are integral to building a true sanitation "value" chain (Cookey et al. 2022; United Nations 2015; US Government 2022). The key indicator categories identified as essential to FSM benchmarking and used in this assessment were: access, containment, collection, transport, treatment, disposal, reuse, social, equity, environment, finance, and management.

We evaluated the pros and cons of indicators within each category through a small-group prioritization exercise with approximately 12 senior researchers from The Aquaya Institute. Specifically, we considered ease of measurement, ease of validation, and reliance on other data sources (e.g., census data). We then assigned a priority level to each KPI, depending on the pros and cons, as well as its application extent and appropriateness for use across multiple countries and contexts. Lastly, we short-listed the highest-priority indicators for each category as recommendations. Prioritization levels were as follows:

- **High**—Used in multiple country programs or integral to global monitoring efforts, represents a critical FSM service aspect, and has few measurement drawbacks;
- **Medium**—Used in one or more country programs, with a relatively even split between pros and cons; and
- **Low**—Limited use in practice, not critical to FSM service, or has limitations that would challenge consistent use across contexts.

2.4 LIMITATIONS

The study methods had some limitations. The literature review methods we employed have some drawbacks relative to more intensive systematic reviews (e.g., comprehensive coverage, reproducibility, rigorous study quality appraisal) but were more time-sensitive and appropriate for gathering information on nascent topics. In addition, due to lags in publishing times and variable reporting formats and languages, we assumed it did not capture all nuance of recent activity on FSM benchmarking and leaned on the case studies to further document experiences. We conducted interviews over a six-week window, which took place during the summer holiday period for some regions and potentially affected the response rate. Some types of organizations (e.g., national governments, private utilities) and locations generally showed less willingness or ability to engage with the research, in which case we aimed to approach multiple people, use multiple outreach methods, or follow up more than twice to achieve adequate data coverage and variety. Recall bias may have affected all of the interviewees, especially when reporting on distant past experiences.

Some of the challenges faced in discussing this topic with interviewees may have stemmed from uneven power dynamics and/or non-standardized terminology. For instance, we encountered dissonance

between local sanitation improvement priorities and the research questions' focus on benchmarking. This may reflect a colonialist legacy of outside donors promoting Western solutions such as benchmarking, which we address in several sections of the report. Both the literature review and case studies reflected that 2023 represented early days for monitoring FSM service delivery, although case study interviewees could easily grasp the vision for potential benefits from their familiarity with the water sector. Terminology related to FSM has been critiqued for embedding assumptions or stereotypes that isolate solutions for low-income countries, when on-site sanitation occurs globally, including in middle- and high-income countries (Strande et al. 2023). Where possible, we adopted the terminology used by the respondent during interviews. Finally, we did not offer the respondents compensation for their time, which likely limited willingness to participate. However, we aimed to fulfill a request from many respondents to share back the final version of the report to, at a minimum, return the value of the knowledge to the organizations from which it originated.

3.0 OVERVIEW OF WATER AND WASTEWATER BENCHMARKING

3.1 DOMESTIC INITIATIVES

Privatization of water supply in the United Kingdom in 1989 first catalyzed adoption of benchmarking practices among water agencies (Ofwat 2023). Since then, multiple countries have established similar benchmarking initiatives. Because many large water utilities also provide sewerage services, benchmarking evolved to include wastewater. For example, the American Water Works Association (AWWA) began benchmarking sewage service providers in 1995 (American Water Works Association, n.d.; AWWA 2020). Domestic benchmarking initiatives have typically fallen into three categories: regulatory monitoring, utility-led monitoring, and enforcement of performance-based contracts (Appendix C).

Regulatory monitoring occurs when a regulatory government body (national or local) tracks service provider performance against policy requirements (Appendix C, Table C1). For example, the Water Services Regulation Authority for England and Wales was created in 1989 to quantitatively track the effects of utility privatization. It originally focused on financial performance but expanded to include a broader variety of performance indicators (Ofwat 2023; TNA, n.d.; Ofwat 2022). In Kenya, the 2002 Water Act tasked the Water Services Regulatory Board (WASREB) with licensing of water service providers and monitoring of their compliance with "minimum service levels" benchmarks over their license validity period (ESAWAS 2022b; WASREB 2022). In 2009, India's Ministry of Urban Development introduced "service level benchmarks," a framework covering water, sanitation, solid waste management, and storm water drainage. After a pilot, all Indian states adopted the framework and still use it to monitor service performance in nearly 1,800 cities (World Bank 2016; Eales and Ahmad 2010; WSP 2018).

Utility-led monitoring occurs when individual service providers or associations of service providers voluntarily establish and monitor KPIs (Appendix C, Table C1). For example, in the Netherlands, where nine out of the ten water companies are public with municipalities and provinces as shareholders, Vewin, the Dutch national association of water companies, initiated voluntary benchmarking in 1997 to resist privatization of the water sector. By improving transparency on water quality, customer service, environmental impact, and financial performance, benchmarking dissuaded interference with the existing service delivery model (De Goede et al. 2016). Although the benchmarking activity began as utility-led, it became mandatory and regulated by the Ministry of Infrastructure and the Environment in 2012 (Accenture 2013). Indonesia provides another example, where the Association of Water Supply Companies, comprised of the country's more than 400 water service providers, introduced a benchmarking program in 2001 with initial financial support from the World Bank. Since 2003, the association fully integrated benchmarking into its operations, with a dedicated budget and staff to collect, analyze, and summarize performance data (Sharma 2006).

Performance-based contracts entail benchmarking when part of the service providers' revenue depends on achieving contractual performance targets (Appendix C, Table C2). Jakarta and Manila were the first two cities in Asia to award large public-private partnership concessions or contracts for water services (ADB 2022). The Government of Indonesia awarded two 25-year public-private partnership concessions for Jakarta in 1997, while Manila, in the Philippines, also awarded two concessions in 1997. Both concessions linked KPIs to a portion of the government payments to the utility (ADB 2022; Rivera Jr. 2013). Performance-based benchmarking can also apply to organizational staff. To improve individual

performance within the organization, Uganda's NWSC monitors KPIs against targets for each of its managers in charge of specific geographic areas (S. Berg 2020; Mugisha, Berg, and Muhairwe 2007).

3.2 INTERNATIONAL INITIATIVES

Many benchmarking initiatives, supported by international donors or multi-country associations of service providers or regulators, extend beyond national borders (Table 4). In addition to fostering standardized guidance, international benchmarking serves two main goals:

1. Guiding donor investments; and/or
2. Voluntary peer-to-peer comparison, in an industry where peers are often located in other countries (specifically in the case of large national utilities).

Stemming from the United Nations (UN) Millennium Development Goal period, the international Joint Monitoring Programme for Water Supply, Sanitation and Hygiene (JMP) implemented by the World Health Organization (WHO) and UNICEF has reported country, regional, and global estimates of progress on drinking water, sanitation, and hygiene (WASH) since 1990 (WHO UNICEF JMP 2023). It extends over more than 200 countries, areas, and territories, producing both comparable longitudinal data and periodic analyses of special issues.

The World Bank-administered International Benchmarking Network for Water and Sanitation Utilities (New IBNET, formerly IBNET) represents another of the largest international benchmarking initiatives. Growing from participation by 15 utilities when first established in 1995, it captures data from more than 4,400 utilities from 135 countries, including 410 utilities in 51 countries in Africa. As a public resource, IBNET data has helped the World Bank identify priority investment areas and has informed several funding programs (C. van den Berg and Danilenko 2011; Danilenko et al. 2014). In addition to curating the world's largest water utility performance database, IBNET develops benchmarking toolkits and guidance and pioneers data validation protocols (Danilenko et al. 2014).

The 2023 relaunch of the New IBNET recognized that the utility benchmarking data collected over more than 20 years has largely served external development groups (i.e., international donors). Following a consultative process, it now aims to better serve utilities themselves and facilitate internal data use. Mechanisms include portals to enter, share, and use metric benchmarking data, as well as to connect with and learn from other utilities (i.e., conduct process benchmarking among self-determined peer learning groups). Built-in benchmarking data visualizations facilitate comparisons over time between organizations and between indicators, revealing insights that help service providers self-assess and work toward performance improvement. The system also facilitates centralized feedback and user data correction. Lessons learned during the relaunch design indicated that users preferred fewer KPIs and selected 15 total for both water and wastewater.

At the regional level, the Eastern and Southern Africa Water and Sanitation Regulators Association (ESAWAS) has connected a network of water supply and sanitation regulators since 2009. It has contributed to peer review and cross-country standardization of minimum service levels. The association developed a regional benchmarking framework in 2015 by harmonizing KPIs and benchmarks used among different member regulators (ESAWAS 2021). As of 2023, ESAWAS includes 12 agencies from Mozambique, Kenya, Rwanda, Tanzania, Zambia, Lesotho, Zanzibar, Malawi, Burundi, Uganda, South Africa, and Angola.

Table 4: Examples of international water and wastewater benchmarking initiatives, sorted in chronological order by start date

Benchmarking Initiative	Start Date	Supporting Entity	Geographic Areas
JMP	1990	WHO, UNICEF	Global (234 countries, areas, and territories)
Utility Benchmarking Program (AWWA 2020)	1995	AWWA	US and two Canadian provinces
IBNET (Danilenko et al. 2014)	1995	World Bank	4,400 utilities in 135 countries
Benchmarking and Performance Assessment Specialist Group (Cabrera et al. 2011)	2000	International Water Association (IWA)	Global
Association of Regulators of Water and Sanitation of the Americas (Corton and Molinari 2007)	2003	Public-Private Infrastructure Advisory Facility[1]	16 countries in Latin America plus Belize
European Benchmarking Co-operation (EBC n.d.)	2007	Voluntary utilities in participating countries	Mostly Europe but also United Arab Emirates, US (1 provider), and Russia (1 provider)
AquaRating (IWA 2021)	2008	Inter-American Development Bank and IWA	29 countries
Global Analysis and Assessment of Sanitation and Drinking-Water (GLAAS, n.d.)	2008	UN-Water, WHO	Global (121 countries and territories)
Pacific Water and Wastewater Association (PWWA 2023)	2010	The Pacific Region Infrastructure Facility[2]	24 water utilities in 14 countries
TrackFin (for financial data) (WHO 2021a)	2012	WHO, UN-Water (GLAAS initiative[3])	Global (21 countries)
ESAWAS 2021	2014	Member water supply and sanitation regulators	10 countries in Africa
Sanitation and Wastewater Atlas of Africa (African Development Bank, UN Environment Programme, and GRID-Arendal 2020)	2020	African Development Bank, UN Environment Programme, GRID-Arendal	Africa (54 countries)

[1] Housed at the World Bank and supported by multiple bilateral and multilateral donors
[2] Housed at the Asian Development Bank (ADB) and supported by multiple bilateral and multilateral donors
[3] GLAAS

4.0 OVERVIEW OF FECAL SLUDGE MANAGEMENT BENCHMARKING

4.1 INTERNATIONAL INITIATIVES

Internationally, since 2014, the Shit-Flow Diagram Promotion Initiative by SuSanA (funded by the Gates Foundation) has been estimating the percentage of unsafely managed fecal waste for 241 cities worldwide (Sustainable Sanitation Alliance 2023). Because it relies on a standardized approach to estimate this city-level performance indicator, in some ways it prefaced international FSM benchmarking. The largely voluntary approach has raised awareness among politicians and decision-makers to advocate for improved FSM but does not typically facilitate repeated or ongoing performance monitoring. In combination with the City Service Delivery Assessment tool, shit-flow diagrams help authorities to assess the baseline situation and diagnose the causes of inadequate FSM (Inclusive Sanitation in Practice n.d.).

In addition, since 2015, the JMP sanitation service ladder has offered a holistic international indicator of sanitation access and safe excreta management, including both on-site and off-site treatment options (WHO UNICEF JMP 2023). In contrast to the prior "improved/unimproved facility" classification that only considered containment, the new definition takes safety over the full sanitation service chain into consideration. The JMP reassesses sanitation indicators every few years going back to 1990.

The JMP and ESAWAS began a workshop series in December 2023 to set priorities for global and local indicators and methods for monitoring on-site sanitation, involving partners from Nepal, Kenya, and Indonesia. ESAWAS has selected several KPIs recommended for monitoring on-site sanitation and developed data collection templates (ESAWAS 2022a). ESAWAS intends to begin reporting on these in their annual regional benchmarking reports as data from member organizations becomes available. The New IBNET initiated a pilot of select FSM indicators chosen in consultation with the CWIS Initiative around the same period, requesting evaluation by several participating service providers.

4.2 DOMESTIC INITIATIVES

Domestic FSM benchmarking initiatives remain limited, reflecting the earlier stage of FSM policies relative to drinking water and sewerage. Institutional and regulatory frameworks governing non-sewered sanitation remain nascent in many places. The 2018/2019 ESAWAS report noted, "the major challenge to improving non-sewered sanitation service delivery in the member countries is the absence of a regulatory framework to address the full value chain of on-site sanitation" (ESAWAS 2019). Still, several locations have made clear steps toward national FSM benchmarking, including Bangladesh, Kenya, Rwanda, Tanzania, and Zambia. Sub-national activities spurred FSM benchmarking in some other locations, including India, Uganda, Indonesia, the Philippines, and Haiti. Notably, international donor activities have influenced FSM benchmarking activities in almost all locations, while other key actors have included city governments, national regulators, NGOs, and public research institutions (Table 1).

National regulation has encouraged FSM benchmarking activities in several cases. In both Rwanda and Tanzania, national regulators issued guidelines for on-site sanitation, establishing KPIs that cover the entire sanitation service chain (EWURA 2020; Tsinda 2020; Nzitonda 2022). In Kenya and Zambia, regulators have taken steps to formalize responsibilities for on-site sanitation in national policies, although KPIs capturing the entire sanitation service chain remain under development. In Zambia, the regulator adopted national FSM standards in 2018, but the only documented FSM KPI relates to percentage of sanitation coverage (NWASCO 2021; 2018). In Kenya, where enactment of the

national policy is pending, the regulator recommends future collaborative development of additional indicators for on-site sanitation beyond JMP sanitation access data (WASREB 2022; Kenya Ministry of Water, Sanitation, and Irrigation 2021; Kenya Ministry of Water, Sanitation, and Irrigation and African Population and Health Research Center 2022).

India and Bangladesh have seen FSM benchmarking activated by a combination of national policies and local support from non-governmental partners. In India, the CWAS under CEPT University's Research and Development Foundation (with funding from the Gates Foundation) was established to serve local governments in 2009 when they launched the Performance Assessment System (PAS) for city-level water and sewered sanitation services in two states (Gujarat and Maharashtra). The country's Ministry of Urban Development recognized the alignment with its Service-Level Benchmarks (Table C1) and endorsed the PAS in 2014 (CEPT University 2018). The Center then added five FSM indicators capturing key themes along the sanitation service chain (toilet coverage, collection efficiency, treatment capacity, treatment quality, and reuse) to the PAS in 2014 (CEPT University 2015). It was challenged in practice by a paucity of verification methods for on-site sanitation data collection, as opposed to estimates (CEPT University 2015), although five states reported data from 2016 to 2021 (CEPT University 2015; CWAS 2021).

Bangladesh introduced a regulatory framework for FSM in 2017 (Ministry of Local Government, Rural Development and Co-operatives 2017) and complementary national action plans in 2020–2021 (Ministry of Local Government, Rural Development, and Co-operatives 2021), with assistance from NGOs such as Practical Action (URBAN WASH 2023; Practical Action 2023). Under the Gates Foundation's CWIS initiative, another NGO (SNV) aided two cities (Khulna and Jhenaidah) with an online information system to manage and monitor FSM services (SNV Bangladesh 2020). SNV proposed a comprehensive list of FSM KPIs, covering all steps in the sanitation service chain (SNV Bangladesh 2021a).

Some municipalities, donors, or nongovernmental actors have begun initiating sub-national FSM data collection and performance measurement in the absence of national regulations. External donor activities, such as the city-level CWIS initiative in parts of Africa and South Asia, have influenced actors such as the KCCA, which holds responsibility for providing FSM services. With support of donors, the city authority created a dedicated FSM unit that strengthens the regulation of FSM service providers through development and maintenance of a sanitation geographic information system (GIS) database and service agreements (KCCA 2019; JICA 2022; CWIS 2021a). Across Indonesia, the donor-supported SANIMAS initiative has stimulated non-sewered sanitation development and accompanying evaluation by local governments, community groups, and external parties inclusive of FSM performance indicators (Bulson et al. 2021). In the Philippines, the government launched the 2012 National Sewerage and Septage Management Program to support cities reluctant to combat the health risks posed by fecal matter overflows. However, municipal leadership, including local stakeholder engagement and bottom-up planning, has been integral to encouraging active FSM implementation and progress tracking (Robbins, Strande, and Doczi 2021).

5.0 ENABLING FACTORS

Enabling factors make it possible (or easier) for individuals or groups to undertake and sustain adoption of benchmarking. Case study respondents generally indicated personal, customized motivations for engaging in benchmarking that illustrated a variety of potential enabling factors with some common threads. These underlying structural, social, and resource supports can potentially contribute to FSM benchmarking across different contexts. This section describes the necessary pre-conditions and building blocks that help to initiate FSM benchmarking, as well as the range of supporting institutional arrangements. The next section describes good practices that sustain effective water, wastewater, and FSM benchmarking programs over time.

5.1 PRE-CONDITIONS

The case study analysis highlighted the importance of nesting FSM benchmarking activities, where feasible, within a context where certain conditions have already been met, including:

1. Basic FSM service elements (e.g., infrastructure; regulatory standards; safety standards for waste handling, treatment, and disposal or reuse; tariff structures and subsidies or other funding sources; human resources training);
2. Readiness for continuous improvement programming among on-site sanitation service provision actors; and
3. Effective lines of communication among the responsible parties.

First, the absence of FSM infrastructure (e.g., emptiable toilet facilities, dumping bays, treatment plants) limited the ability to carry out FSM services and therefore evaluate and improve the service performance. In addition, **supporting service elements**, such as sludge quality standards, market research, worker safety standards, and agreed pricing are needed to accurately define and measure a full range of functional benchmarking indicators. In Zambia, for example, standards to confirm the safety of applying sludge reuse products to crops for human consumption (as opposed to landscaping) remained under development, forestalling promotion and tracking of sludge reuse indicators by the national regulator. In practice, it is challenging to disentangle factors supportive of holistic on-site sanitation from factors supportive of benchmarking.

Actor **readiness** (i.e., commitment to implementing an organizational change and confidence in collective abilities) typically applies to the service provider's organizational climate (Weiner 2009). However, it may also encompass the country context external to the service provider (e.g., underlying sociopolitical, financial, and environmental stability), as noted in the Sanitation Social-Ecological System (S-SES) model. For example, a national election transition delayed the onset of a national FSM policy in Kenya. Meanwhile, a paucity of national regulatory agency staffing for sanitation, which has stagnated over about 10 years, has limited formal FSM oversight activities in Haiti. Violence and mistrust in Haiti has further hampered the financial and infrastructure support needed to address the full sanitation service chain (DINEPA, Lasante, and DAI 2022). Organizations distracted by other pressing issues are unlikely to be able to focus their efforts on benchmarking.

Third, the case studies examined multiple possible service provision arrangements, all involving multiple stakeholders. To ensure benchmarking data can be used to improve practice and policy, regular multidirectional **communication pathways** should exist among the actors involved in collecting sludge, carrying out treatment and disposal, allocating funds, and overseeing service provision (Bartram and Setty 2021). Challenges to communication observed in the case studies included a lack of feedback

mechanisms, liability concerns, having one actor with greater readiness than others, or misaligned mandates and budgets. Over time, on-site sanitation responsibilities have tended to follow after sewered sanitation service responsibilities, which followed after water delivery responsibilities, contributing to potentially inconsistent, insufficient, or redundant institutional arrangements. In Bangladesh, for example, separate local government departments are responsible for implementation and finance, precluding assessments of FSM financial performance (SNV Bangladesh 2021b). Further, respondents from Bangladesh noted that the municipalities tend to look to central government investment projects rather than plan from their own resources, and larger cities may have multiple corporations or authorities in charge of sanitation that fail to cooperate. Another case highlighted the need for FSM service providers to pursue duplicative engagement efforts, due to overlapping mandates among Kenyan ministries. In contrast, a Zambian utility noted the national regulator's leadership and 2018 FSM Framework greatly helped to clarify the roles and responsibilities of various actors (NWASCO 2021, 2022).

5.2 BUILDING BLOCKS

Building blocks represent active steps toward applied benchmarking. Four key building blocks that can launch FSM benchmarking activities at the national level include: a favorable regulatory environment, capable data champions within participating organizations, a business model supported by an active customer base, and financial support tied to the benchmarking activity (Figure 1). While these stem from the case studies, any of the building blocks might serve as an initial entry point to facilitate the introduction of FSM benchmarking programming in other locations. Ideally, all four work in conjunction to support benchmarking program effectiveness over time.

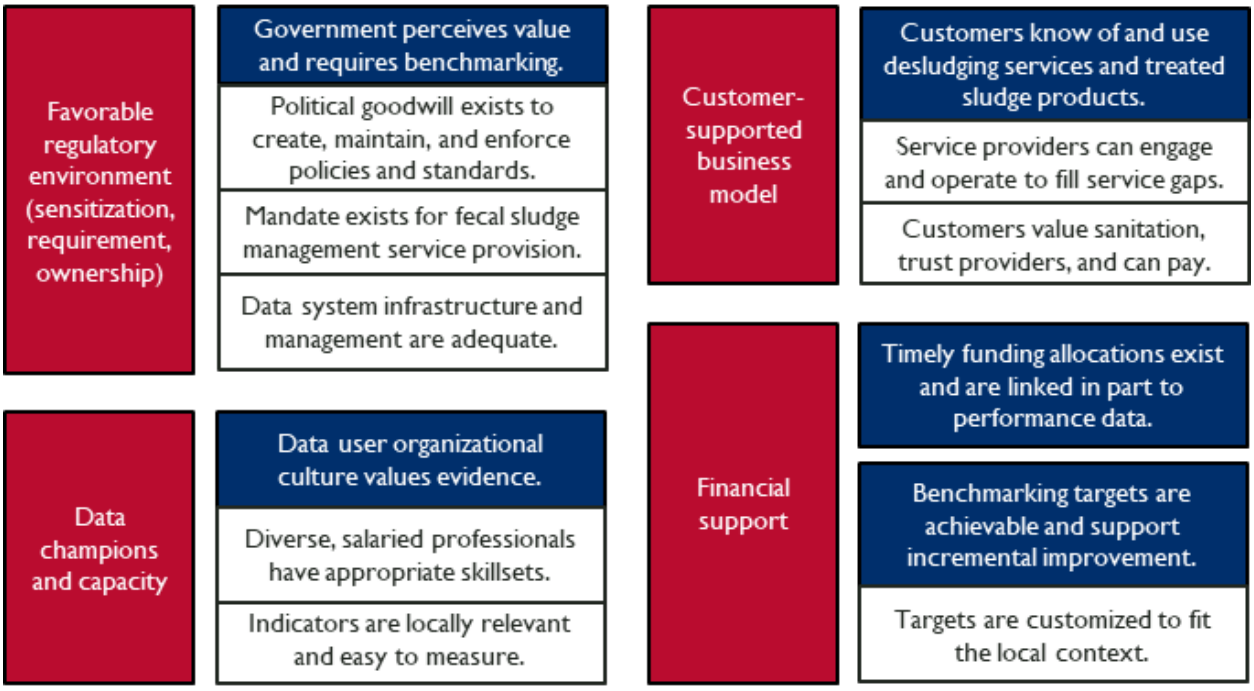

Figure 1: Four primary building blocks for FSM benchmarking observed across the combined case study sample

The **regulatory environment** building block captures a spectrum of progress from awareness to active ownership and enforcement. This mirrors a concept in the WHO's Global Benchmarking Tool for evaluation of national regulatory systems of medical products, which uses a "maturity level" approach

adapted from ISO 9004 (organizational quality management) on a scale from existence of some regulatory system elements to operation at an advanced level of performance and continuous improvement (WHO 2021b). In Indonesia, for example, respondents felt the national government had been sensitized, and further steps to hold them accountable could contribute to mobilizing available resources and engaging local governments. Even thorough national regulations or policies could fail in practice or fade in relevance over time if they lack implementation oversight mechanisms.

The **customer support** building block reflects the historical origination of benchmarking as a tool to increase business competitiveness, although it broadly applies to multiple service provision models (e.g., public-private partnerships, social enterprises, NGOs, or public services), as observed in the case studies. To improve service performance, providers must engage with customer demand. Without active customer transactions, much fecal sludge generation and flow (e.g., direct environmental discharge) would remain unmonitored. In Bangladesh, a key priority is to increase demand for household emptying services by raising community awareness of Khulna's new fecal sludge treatment plant, which has not reached its full service capacity. Marketing is also a priority in the Philippines, where customers frequently turned down sludge removal services and the status quo favors informal services or unnoticed environmental overflows (Robbins, Strande, and Doczi 2021). This led to new approaches, such as no-cost scheduled desludging (as long as the property could be accessed) and creation of targeted KPIs for service uptake.

Regarding **data champions and capacity,** a subset of more specialized and engaged practitioners can effectively promote benchmarking. Data use for decision-making is likely only a moderate priority for most professionals (The Aquaya Institute 2022b). Thus, not all staff of FSM service organizations need intensive training to achieve general awareness, cooperation, and support of benchmarking programs. A private emptier in Uganda remarked on widespread acceptance of quantitative data reporting among their peers, even though they might place greater value on qualitative experiences individually. Voluntary data submitters (e.g., contributors to the SuSanA or early adopters of IBNET) might be considered data advocates. In Haiti, a nongovernmental service provider that values evidence-informed decision-making uses KPIs for many purposes from employee reviews to setting strategic targets with their board. They also report performance data to the national sanitation authority, although it is not formally required, and have advocated for embedding such reporting into future service provider oversight.

FSM benchmarking efforts can benefit from **financial support** or advantages that depend on actors reporting KPI data or meeting specific target goals tailored to their local context. Financial pressure could be exerted in two directions: for instance, service providers that receive dedicated performance-based funding may be incentivized to comply with benchmarking programs, while others might achieve better financial outcomes by not cooperating with formal reporting requirements. Still, the other building blocks (strong regulatory enforcement, customer demand, or champions) could push actors to view participation in benchmarking systems as advantageous to their bottom line and allocate staff time or internal performance incentives to support it.

Multiple case studies illustrated the key role of donors in influencing FSM benchmarking activities and the challenges of transitioning to independently funded benchmarking arrangements (e.g., Narsapur in India; CWIS 2021c), such that ongoing benchmarking efforts clearly align with the value gained. The city authority in Kampala reflected on the risk of continuing to spend heavily to collect KPI data (previously funded by donor initiatives) that is not used to inform planning or decision-making, potentially overstretching already limited internal resources (Box 1). A private emptying company in Uganda closely intertwined benchmarking with their business operations and profit potential. Respondents from a social enterprise likewise saw performance data reporting as critical to continued donor investments. Less clarity surrounded the existing links between benchmarking and financial opportunities in the public

sector; one respondent from Zambia indicated that regulatory incentives issued to utilities might trickle down to staff in the form of personal incentives.

Box 1: Cooperation to extend FSM benchmarking that began as a donor-supported activity in Kampala, Uganda

Donor funding has supported FSM capacity strengthening in Kampala since 2017, especially through the Gates Foundation's CWIS Initiative and its corresponding monitoring and evaluation approach. The majority of residents in Kampala, Uganda, use on-site sanitation, and the city service provider has made substantive progress in serving low-income areas. Because of the large number of CWIS-related KPIs and their sometimes-technical nature, the KCCA is actively working to streamline and optimize the number of KPIs used for ongoing monitoring.

Transitioning to a more permanent model will require coordination to institutionalize financial support for FSM and FSM benchmarking at the city and national level. Ugandan city governments have the mandate to manage fecal sludge under existing laws but operate without much centralized financial or technical support outside of Kampala. Within the capital, the KCCA FSM unit should oversee sludge collection and transport but has only limited funding to maintain trucks. In contrast, the NWSC retains responsibility for sludge treatment and a larger budget allocation from sewer customer revenue. Further, several ministries (Water and Environment, Health, Education, and Government) must cooperate to provide consistent FSM policy guidance. Kampala City and the GIZ have supported professionalization among independent private emptiers, who underwent training to form consolidated businesses that formally report benchmarking data. These businesses can participate in national and regional professional emptier associations (e.g., Association of Uganda Emptiers, Pan-African Association of Sanitation Actors), which support peer learning and process benchmarking.

5.3 INSTITUTIONAL ARRANGEMENTS

Particular approaches to initiating or establishing FSM benchmarking systems differ among the different institutional or governance contexts found around the globe. Still, common themes related to relationships, shared benchmarking goals, and stakeholder inclusion played a role in facilitating benchmarking across the range of institutional arrangements seen in the case studies.

The case studies reflected the feasibility of promoting FSM benchmarking given diverse institutional arrangements. In Bangladesh, for example, the NGO SNV provided legal support to create contracts between FSM service providers and municipalities, easing and incentivizing formal cooperation between the two actor types. In the Philippines, a public-private partnership split utility service provision responsibilities for different areas of the city between Manila Water and Maynilad, who report to the same metropolitan regulator. This allows the regulator to easily benchmark the two groups against each other to understand the relative feasibility of performance improvements, enhancing comparability and competition while disincentivizing complacency. Initiating benchmarking as part of municipal or city-level demonstration projects, as seen in Lusaka and Kampala, could provide a starting model to further FSM efforts at the national scale (Box 2).

Box 2: Zambia's capital serves as a demonstration case for FSM benchmarking expansion

The World Bank and CWIS Initiative sponsored by the Gates Foundation has supported FSM benchmarking activity since approximately 2012 in Lusaka, Zambia. The national regulator NWASCO received praise for its 2018 reform that explicitly added on-site sanitation in urban utility mandates. To operationalize the new service mandate and foster continued benchmarking, LWS rolled out performance-based contracts with multiple private emptiers. A multi-stakeholder national technical working group made recommendations to LWSC about which KPIs to use. Zambia's National Assembly and Ministry of Water Development and Sanitation have engaged substantively with Lusaka's FSM benchmarking reports, sometimes calling for special sessions to discuss questions. Implementing partners regularly refer to Lusaka when designing new FSM and FSM monitoring programs for other urban centers. Still, attention to factors that support FSM at the national level, such as establishing sustainable public funding for FSM (beyond the sanitation surcharge on water bills traditionally used to support sewerage), tariffs, subsidies, household sanitation facility standards, and licensing, would help to make both FSM operations and benchmarking data collection smoother.

Because FSM service delivery models varied widely and no real-world setting will likely have one "ideal" set of institutional arrangements, **benchmarking can be used to achieve other goals, such as issue visibility or advocacy**, until other necessary FSM components are in place to progress toward service improvement. For example, shit-flow diagrams have effectively illustrated a lack of sanitation progress for advocacy purposes (Dubois 2023). In the Haiti case study, voluntary service provider reporting on KPIs has preceded other building blocks, such as regulatory mandates, while the sociopolitical and environmental context poses extreme challenges and the vision of widespread privately operated FSM systems remains out of reach (DINEPA, Lasante, and DAI 2022).

Respondents generally expected guidance and feedback to come from larger-scale institutions (e.g., public utilities, legislators, national regulators, regional professional networks). If they did not receive a response, they tended to assume no news was good news or tried to gauge effort among their peers. For instance, an emptying company in Kampala did not know how data might be used at higher levels but assumed they could continue operating as usual if they did not hear about problems. An NGO noted the lack of clear guidance on establishing subsidies in Zambia, resulting in a tendency to copy the efforts of other peer groups operating in the same country instead.

The case study analysis illustrated that missing contributions from any FSM actor type would have negative implications for benchmarking.

In particular, **research and learning institutions** had clear roles regarding fecal sludge treatment and reuse safety, standard setting, and data validation. These activities may be formalized within institutions that directly serve public agencies, as in India's CWAS or Bangladesh, where an FSM support unit was installed at ITN-BUET (Al-Muyeed, Mukherjee, and Hassan, n.d.). Alternatively, independent NGOs, private auditors, or educational institutions may lead applied research within a given country. A combination of different types of in-country and external research, think-tank, or media organizations could provide checks and balances that optimize innovation, data validation, and research application.

Private sector involvement can likewise strengthen service delivery options, including in hard-to-reach low-income areas such as informal settlements. A private emptier in Uganda had the flexibility to provide free services to assist the elderly and people with disabilities as part of their corporate social responsibility agenda. A benchmarking activity with South Africa's eThekwini municipality described a "triple helix" of cooperation between universities, the private sector, and government (Senkwe and

Niwagaba 2016). Another example from Odisha, India, highlighted formation of private FSM operations as a vehicle to grow social inclusivity by extending dignified work and income opportunities to marginalized communities, such as survivors of domestic violence or transgender individuals (Basyal 2023).

We observed little **local consumer participation** to support benchmarking programs. In Bangladesh, the Local Government Act mandated public participation through town- and ward-level committees. However, effectiveness in practice largely depended on municipal leadership (mayor and council). One case took the "customer survey" approach to benchmarking (IBNET 2021a) as a way to supplement metric benchmarking: In Kampala City, the Weyonje app helped to track customer feedback as part of a larger CWIS community engagement campaign, so the public could easily communicate concerns to emptiers and the city government. Kampala City had historically relied on global (CWIS) or regional (ESAWAS) actors to provide benchmarking guidance, but following these experiences, they moved toward involving local stakeholders in customizing indicators. To level the playing field, donors or regulators could use their relatively outsized influence to require public stakeholder participation in benchmarking processes.

6.0 GOOD PRACTICES IN BENCHMARKING

Following known good practices can help to sustain benchmarking activities over time and foster beneficial outcomes. Good practices that emerged from this study revolved around three themes: (1) data, (2) incentives, and (3) dissemination. These themes provided common denominators across all forms of benchmarking (international, domestic, regulatory, utility-led, and performance-based contracts), with applicability to water, sewered sanitation, on-site sanitation, and solid waste management.

6.1 DATA MANAGEMENT

Effective benchmarking relies on good quality and trustworthy data. Core recommendations emerged from the water and wastewater literature, while early examples in the case studies highlighted more complex data management aspects of FSM benchmarking:

Focus on a short list of verifiable and useful indicators. "Having an exceedingly long list of KPIs will make monitoring unwieldy and the linkages prone to contest by both parties" (ADB 2022). Continuously collecting and verifying performance data requires intensive resources for data collection and verification (McDonald 2016; CEPT University 2015); therefore, limiting the number of KPIs makes benchmarking more approachable, particularly for small utilities (Thiadens and Betham 2013). A small number of simple KPIs, on the order of five, eases communication and interpretation by local government officials (Still and Balfour 2006; Sanjeevi and Shahabudeen 2015). KPIs should be selected to maximize usefulness for addressing management questions. Understanding whether the data might inform a specific course of action can assist indicator prioritization and presentation.

For FSM benchmarking, capturing the sanitation service chain holistically may require multiple indicators reflecting the efforts of diverse service providers, making it difficult to limit the number of KPIs. For example, the FSM regulatory frameworks in Rwanda and Tanzania involve 13–17 KPIs (RURA 2021; Nzitonda 2022; EWURA 2020). In India, experts have proposed up to 16 KPIs (Jayathilake et al. 2019; Velidandla et al. 2020), while the CWIS Initiative includes more than 30 indicator areas ("CWIS Measurement Note" n.d.). The process of reaching consensus on a small set of indicators may take up to several years, as demonstrated in the solid waste field (AECOM 2012).

Define shared indicators well and ensure they apply widely across contexts. For example, staff productivity may be challenging to measure and compare across different contexts. In addition, studies have debated the best definition for non-revenue water, potentially obscuring comparisons across service providers (AWWA 2019). In the FSM case study in Uganda, respondents noted that poor comparability of shared benchmarking data mainly stemmed from differences in organizations' definitions. This issue further pointed to a lack of accurate FSM records, maps, or inventories to interpret what exactly the data represented. For example, the citywide ratio of treated to untreated sludge might not distinguish where exactly the sludge was collected, potentially obscuring disproportional impacts on low-income or informal urban neighborhoods.

Apply multiple measures to improve data quality, as benchmarking largely relies on self-reported data. Water and wastewater benchmarking examples have traditionally used one or more feedback steps to ensure data quality. For example, the New IBNET embeds quality controls in its data collection tool, which automatically flags when data fall outside expected ranges or appears unrealistic given historical trends. The IWA and IBNET have both historically attributed reliability scores to KPI values, based on the strength of supporting evidence and documentation (IBNET 2021b; Krause et al.

2018). Similarly, the Tanzanian regulator assigns confidence levels to KPI data submitted by individual utilities, which factor into utility managers' performance records (S. Berg 2020).

The FSM benchmarking case studies suggested easing data validation by combining multiple feedback methods. Interviewees from Zambia and other locations did not consider existing data spot checks as providing adequate coverage to confirm data quality and validity because oversight agencies use them at infrequent intervals. Typically, oversight agencies' staff time, availability, and budgets severely limit these efforts. Respondents from India in particular noted third-party data verification as a key gap affecting benchmarking efforts. To overcome this issue, robust automated data checks could be built into benchmarking data systems to avoid human errors and tampering where possible, with supplemental human verification measures. Following standard quality assurance protocols, including automated time- and user-stamped records, would make it easier to maintain transparent metadata.

Anticipate the personnel and equipment requirements. Collecting, verifying, and aggregating performance data may require dedicated field and office skills training, equipment, and staff. Some FSM operators may view data collection as an imposition or find that inputting digital data doubles their workload if they usually keep paper records. They may also experience pressure to avoid tracking improper procedures (e.g., those that shortcut taxation, worker safety, or environmental protections). Desludging operators in India reported reluctance to submit digitized data because some avoid bringing their phone to work to prevent soiling or damage (Athena Infonomics 2020). In addition, operators may have internet connectivity issues or find mobile applications difficult to use. Lessons from Lusaka, Zambia, and Maputo, Mozambique, substantiated this finding where an NGO piloted a data tracking mobile application for FSM operators (called Pula) only to learn that the intended users found the application too complex (The Aquaya Institute 2021).

Anticipation of the added labor, resource, and supporting data requirements that go into benchmarking programs can smooth startup. Kenyan stakeholders welcomed personnel training and outfitting of regional data hubs with computing equipment to support benchmarking scale-up. Kampala City reported shifting from resource-intensive manual data collection to paperless digital solutions and integrating GIS maps to assist data verification. Because household surveys would be cost-prohibitive in many cities, it may be possible to leverage information collected during service visits through property tax assessment forms (CEPT University 2015) or other existing (e.g., census) records.

Additional strategies may be needed to encourage data submission by lower-capacity service providers. For water and wastewater, this involved relying on community-based organizations as done in Pakistan to monitor water access and quality (Maqbool, Shahid, and Khan 2022). Another strategy consisted of donor agencies managing KPI data, as the World Bank did in the Danube region (S. Berg 2020). The FSM case studies did not reveal novel strategies specific to ensuring equitable data submission participation, although the larger proportion of informal service providers in low-resource settings recommends incentivizing business formalization as a first step.

Integrate fecal sludge data management with other reporting systems. Although not noted in the literature review, some evidence from the case studies suggested that merging FSM benchmarking efforts with other benchmarking processes could ease startup and labor burdens, while aligning data management practices to programs that have synergistic goals. For example, case study respondents recommended leveraging other data reporting systems (e.g., village health in Uganda, solid waste management in Bangladesh) for long-term sustainability (Aspire to Innovate (a2i) Programme 2023). This could also help connect FSM service providers to appropriate technology support options, such as existing mobile apps or data systems.

6.2 INCENTIVES

The water and wastewater sectors have used financial, reputational, and/or regulatory mechanisms to create incentives (Appendix C, Table C1). The FSM benchmarking case studies demonstrated more complex applications mixing contractual, individual, organizational, financial, and regulatory rewards and penalties to reinforce desired behaviors:

Consider opportunities for direct and indirect financial rewards. Stakeholders can integrate direct financial rewards into performance-based contracts that contractually link KPI target achievement to payments. These rewards may apply to individuals and/or institutions. For example, Uganda's NWSC provides salary bonuses for area managers who achieve KPI targets and penalizes via salary reductions those who do not meet targets (C. van den Berg and Danilenko 2017). In the Philippines, part of the government payments to Manila Water link contractually to KPIs (ADB 2022; Rivera Jr. 2013). In Zambia, the timing of a daily cash incentive (top-up) for sludge delivery to the treatment plant by volume played a critical role for emptiers as part of their performance-based contracts with LWSC as it ensured regular cash flow to pay expenses. Financial disincentives may include direct financial penalties, such as deductions from operator or service provider compensation, as part of performance-based contracts (ADB 2022).

Studies also found that eligibility for additional funding, such as access to grants or loans, incentivized KPI monitoring and performance improvement (GIZ 2019). In Uganda and Tanzania, indirect performance-linked funding takes place in the form of local development grants, whereby fund transfers to local governments depend on demonstrating good financial management (CEPT University 2010). In India, under the Jawaharlal Nehru National Urban Renewal Mission, national and state governments provide capital funds for water supply and sanitation as grants on the condition that local governments recover 100 percent of their operation and maintenance costs through local charges and taxes (Mehta and Mehta 2014). Additionally, states taking part in benchmarking become eligible for grants from the Central Financial Commission (Appendix C, Table C1). Another form of financial incentive links tariff adjustments to KPI requirements (S. Berg 2013), as in the Manila FSM case study or in Zambia where NWASCO relies on KPI data when reviewing tariff adjustment requests (NWASCO 2014).

Regulators can penalize poor performance or lack of participation in benchmarking. The United Kingdom's Water Services Regulation Authority can impose enforcement orders and, for the most egregious low performers, levy financial penalties on service providers (Ofwat 2017). Zambia's NWASCO issues enforcement notices to non-compliant providers and penalties after two notices, requiring submission of performance improvement plans and monthly reports that highlight progress (ESAWAS 2019). In the Uganda case study, regulators could cancel sludge emptier licenses (operational permits) at any time if the entity failed to correctly report benchmarking data.

Reputational rewards represent another important strategy to increase motivation for improving performance. Reputational rewards may entail publishing a ranking of cities or service providers and assigning financial or reputational rewards to high performers. In Zambia, the national regulator presents Chief Executive Officer awards, encouraging and giving visibility to service provider leaders who have achieved good benchmarking results (NWASCO 2021). In Kenya, WASREB designates "top-performing utilities" in a number of categories and provides awards to those service providers on an annual basis (WASREB 2023). India's National City Rating (list of cleanest cities) offers an example of a city-level reputational incentive implemented under the Swachh Bharat Mission (to end open defecation and manage solid waste). It uses data from a nationwide annual cleanliness survey (Swachh Survekshan) conducted by the Ministry of Housing and Urban Affairs since 2016.

While incentives for FSM benchmarking remain in their infancy, similar reputational rewards might help to engender improved work cultures, partnerships, and subsequent financial rewards that raise the overall quality of FSM services. In South Africa, a "brown drop" certification system for non-sewered sanitation systems would help customers verify responsible management (Pillay and Bhagwan n.d.). A service provider in Uganda believed their formal licensing designation helped affirm to customers that the company abides by professional standards and offers a higher level of service than informal providers, aiding customer acquisition.

6.3 PUBLIC DISSEMINATION

Publicly disseminating benchmarking data promotes transparency and accountability for service delivery and can foster competition. However, this requires that data are trustworthy, as described above.

Share data broadly, if the data are reliable. Many benchmarking initiatives involve public data dissemination (e.g., 12 out of 15 countries in Table C1). For example, Brazil, Zambia, and Kenya issue annual online reports with detailed KPI data and a ranking of service providers according to their performance (CEPT University 2010; Franceys, Drabble, and Renouf 2020; ESAWAS 2019). Similarly, international benchmarking initiatives such as IBNET and ESAWAS make data public through online dashboards and reports, although the New IBNET will take measures to de-identify and aggregate data from multiple service providers within the same country. In addition to promoting transparency and accountability, public data dissemination also helps to accomplish the function of benchmarking as a substitute for direct market competition (McDonald 2016).

In practice, effective public data dissemination among different contexts may depend upon a more complex mix of factors such as high-level buy-in, correct and validated data, relatively favorable benchmarking data, engagement incentives, and maintenance of a user-friendly interface. FSM benchmarking reports were not usually highly accessible. Data exchange among key stakeholders in Kampala often happened previously only on request (CWIS 2021a). Some case study respondents (e.g., in the Philippines, India) expressed hesitancy to share benchmarking information, potentially due to confidentiality concerns, liability, or reputational risks. Contacts in other locations (e.g., Zambia) expressed widespread concerns about public utilities potentially manipulating data in a more positive direction due to an obvious disconnect between the data portrayed and the lived experience of residents and NGOs. While complete quality control of benchmarking reports may pose a challenge (as noted above), following good practices and documenting the methods used can eventually increase transparency and trustworthiness.

6.4 SCALE OF BENCHMARKING

Multiple actors can carry out benchmarking of themselves and/or others at different spatial scales. To promote competition and learning, benchmarking often compares KPI values within peer groupings. For example, Uganda's NWSC applies domestic-level benchmarking through performance contracts with each of its area managers and takes part in regional inter-country benchmarking by ESAWAS. The FSM case studies revealed a wide range of possible motivations and applications of benchmarking, ranging from service providers who deal closely with customers tracking their own indicators for internal business purposes to city-level, national, or global actors comparing service performance over larger spatial scales. Table 5 lists some advantages and limitations of different scales of benchmarking.

Table 5: Advantages and limitations of different scales of benchmarking

	Advantages	Limitations
International (Global)	• May help scale innovative approaches, such as the use of pro-poor, gender, or climate-related KPIs • Helps to drive achievement of global development goals	• Challenges comparability in KPIs as utilities may be reporting against different country-specific regulatory standards and priorities • Limits incentives to reputational rewards because international institutions do not have any regulatory power over utilities and cannot impose sanctions or provide financial incentives
International (Regional)	• Provides a comparison for national utilities (e.g., Uganda's NWSC or Ghana Water Limited) that have a monopoly and therefore do not have a comparison point domestically • Expands options for "model" systems used for process benchmarking that are both contextually relevant and achievable	• Requires individual countries to have sufficient data management capacity and infrastructure (e.g., Burundi as noted in ESAWAS 2021) • May differ from indicators used globally or in other regions (e.g., if some countries respond more quickly or slowly to emerging issues)
Domestic (National)	• Links more easily to regulatory responses and financial incentives (e.g., payments under performance-based contracts, eligibility for central government funding, approval for tariff adjustments) • Promotes peer learning among cities within the same country, which is subject to the same economic, institutional, and regulatory environment	• Depending on the country, may involve broader or narrower perspectives relative to global industry standards (e.g., sociopolitical feasibility of considering gender equity and social inclusion indicators) • National agencies reporting data may have limited oversight and accountability (e.g., from lawmakers, media groups, academics, civil society, or regional networks)
City Utility	• Captures within-country variability • Creates opportunities for peer-to-peer process comparisons • May attract opportunities for additional financial support	• May neglect rural, unincorporated, informal, or peri-urban areas, potentially perpetuating domestic inequities • Often requires cooperation among multiple service providers
Service Area Provider (e.g., private vendor)	• Provides higher resolution data • Promotes coordination among service providers to address service gaps • Offers targeted information to improve business practices (and profitability or financial sustainability) • Can leverage customer interactions to collect data	• May neglect low-income areas or informal (e.g., manual) services • May result in punishment of informal service providers, further limiting service improvement options in marginalized communities

Due to the wide range and scale of benchmarking applications, the KPIs used for different purposes may also vary (Figure 2). The number of KPIs typically reduces as the spatial scale increases and stakeholders aggregate or sub-select indicators for broader comparison. In the case studies, we observed that some indicators may be selected for internal management uses rather than external sharing or comparison. For example, service providers in Haiti and Uganda collected some indicators mainly for personnel or financial management while setting strategic targets or more widely sharing other indicators as part of strategic year-over-year or peer benchmarking. Thus, standardized and shared versus locally adapted indicators may serve unique purposes and hold different value for organizations.

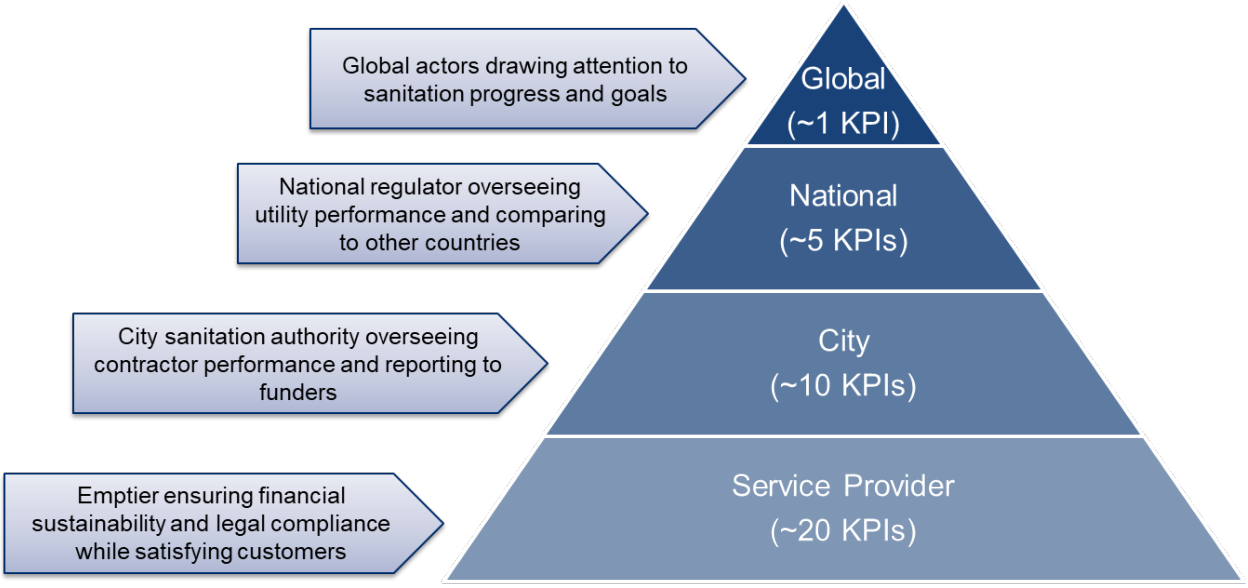

Figure 2: Example tailoring the extent of FSM benchmarking activity and indicative number of KPIs to stakeholder purposes at different spatial scales

7.0 KEY PERFORMANCE INDICATORS

Benchmarking primarily collects, aggregates, and validates data on KPIs that characterize service delivery and management. Examples of KPIs used in the water and wastewater sector appear in Appendix D (Table D1), covering operational aspects such as service coverage and reliability, financial aspects such as operating cost recovery, and organizational aspects such as staff productivity and gender equity.

7.1 FECAL SLUDGE MANAGEMENT INDICATORS

Compared with water and wastewater benchmarking, it is more difficult to keep the list of FSM KPIs short because the on-site sanitation service chain involves multiple steps, each typically under the responsibility of different entities (e.g., households, desludging operators, utilities operating treatment plants). FSM benchmarking thus requires collecting information from a patchwork of entities, often including many small operators, compared to a single service provider for water and wastewater benchmarking. The case study experiences and New IBNET relaunch (15 indicators for water and wastewater) suggested selecting a manageable number of shared KPIs on the order of 10, in contrast with the higher numbers (up to 40) previously trialed by CWIS, although the precise numbers may vary upward or downward depending on the spatial scale and actor purposes (Figure 2).

7.1.1 PERFORMANCE VERSUS MANAGEMENT INDICATORS

Some benchmarking initiatives, including the New IBNET, have moved toward segregating service delivery performance indicators from management practices and financial health indicators. Management practices might include regulatory conditions, procedures, workforce characteristics, and comparisons between income and expenditures. Because the New IBNET will record self-assessed management practices (a form of process benchmarking) confidentially to encourage honesty, utility users can view visualizations correlating their own management practices with service delivery performance indicators. Once multiple utilities report data, it can display publicly in aggregate. Pilot indicators specific to on-site sanitation (Appendix D, Table D2) are being tested and not yet part of the live dashboard. The Philippines FSM benchmarking case study took a similar approach to separating management practices and financials from other service delivery indicators.

7.1.2 RECOMMENDED KEY PERFORMANCE INDICATORS

The literature review initially identified a small number (n = 10) of KPIs commonly used in FSM benchmarking. The case study KPI collation and prioritization exercise (Appendix D) provided more extensive information on multiple options tailored to global, national, and local contexts. We assessed the full set of FSM KPIs (n = 98) via a more extensive review of the case study information and additional considerations such as the need for other data sources, existing extent of use, potential alternative indicators, benefits, and measurement challenges (Appendix D, Table D2). The table excludes indicators typically used for baseline on-site sanitation situation analysis, focusing instead on those that could potentially be valuable for repeated (e.g., annual) monitoring.

By comparing the options within each category, we proposed a preliminary set of recommended FSM monitoring KPIs (Table 6), which prioritize city-level indicators for global comparison. Several have more extensive past use while we included others to advocate for greater visibility moving forward (also see priority ranking definitions in section 2.3 or Appendix D, Table D2). All of the recommended indicators have been proposed or used in various contexts but require continued testing with operators and regulators to confirm their feasibility, interoperability, and comparative value across settings. Using these indicators consistently among locations and at different spatial scales could enhance ongoing data

compatibility and comparisons while drawing greater attention to underrecognized components of the true sanitation value chain (e.g., social responsibility, equity, and environmental impact).

Table 6: Prioritized list of benchmarking KPIs recommended for evaluating FSM services at the city scale across international locations (source references in Appendix D, Table D2)

No.	Category	KPI	Measurement Guidance or Definition	Used Since	Extent of Use
1	Access	Percentage of population with access to safely managed sanitation*	Summary of population coverage, meeting multiple criteria across sanitation service chain (use of an improved sanitation facility that is not shared with other households and where excreta is safely disposed in situ or excreta is removed and treated offsite)	2015	Global (JMP), including Kenya, Rwanda; Regional (CWIS cities, African Sanitation Policy Guidelines)
2	Equity	Access to sanitation services among vulnerable groups	Percentage of population with access to on-site sanitation services disaggregated by income, sex, age, race, ethnicity, migratory status, disability, and/or geographic location as appropriate. For example, comparing poorest versus richest wealth quintile, informal settlements versus formal, or disadvantaged groups versus the general population.	2021	Proposed globally (UN Special Rapporteur on Water and Sanitation); Regional (African Sanitation Policy Guidelines)
3	Containment	Coverage of on-site sanitation facilities	Percentage of the population using on-site sanitation (e.g., pit latrines, container toilets, or septic tanks rather than shared public toilets, no toilets, or sewered toilets) out of the total population	2015	Global (aligns with shit-flow diagrams and JMP and New IBNET sanitation indicator), including India, Zambia, and Uganda
4	Collection	Total collection volume	Volume of fecal sludge collected (sum of daily volumes)	2016–18	Uganda, Rwanda; proposed in Zambia
5	Treatment	Treatment ratio (volume treated: collected)	Volume of properly treated sludge (both liquids and solids) as a percentage of the volume collected	2015	Global (aligns with shit-flow diagrams); Regional (ESAWAS), including Tanzania
6	Disposal/ Reuse	Percentage of fecal sludge safely managed	Volume of safely disposed or reused sludge out of the volume produced; considers a combination of service delivery outcomes from on-site sanitation (e.g., sludge treated on-site, transported and treated offsite, or safely disposed or reused) such that the hazard level and population exposure result in a low public health risk	2015	Global (aligns with shit-flow diagrams); Regional (African Sanitation Policy Guidelines)
7	Social	Worker safety procedure adherence	Percentage of services delivered following occupational health and safety guidelines, out of all services delivered	2019	Piloting globally (New IBNET on-site sanitation indicator); Regional (CWIS cities), including Bangladesh

No.	Category	KPI	Measurement Guidance or Definition	Used Since	Extent of Use
8	Finance	Operating cost recovery	Percentage of operational costs recovered through customer tariffs, government cross-subsidies, and revenue-generating activities for FSM	2019	Regional (CWIS cities for treatment plants only); Bangladesh, Zambia, Haiti; Philippines (through revenue only)
9	Management	Percentage of desludging providers operating with a formal license	Number of licensed, contracted, or regulated providers (including those involved in constructing toilet facilities, emptying, transport, and treatment services) divided by estimated total number of providers	2021	Piloting globally (New IBNET); Regional (African Sanitation Policy Guidelines)

*Excreta separated from human contact via improved sanitation facilities not shared with other households (contained); treated and disposed of on-site; or stored, collected, and transported off-site for treatment (WHO UNICEF JMP 2023).

The access indicator reflects the data needed to support global JMP progress monitoring on sanitation development, whether through on-site or centralized mechanisms (WHO UNICEF JMP 2023). The containment, collection, treatment, and disposal/reuse indicators all contribute to the overarching definition of safely managed sanitation access and are more specifically separated here for the portion of the population using on-site sanitation. These indicators also track with steps in the traditional sanitation service chain and shit-flow diagrams (SFD Promotion Initiative 2018). The containment indicator emphasizes the necessity of knowing the nature of household-level sanitation facilities, which can be used to estimate generated volumes of sludge. The collection indicator recognizes variability in the number of people using each type of facility and rate of use, focusing instead on the easier-to-measure collection volume. The treatment and disposal/reuse indicators signify the extent of environmental and public health protection via pathogen containment or reduction, although the degree of treatment needed and appropriateness of disposal methods may differ among locations. In most cases, these indicators of formally managed FSM services presumably lead to better population health outcomes than alternatives such as direct environmental discharge or open defecation (SFD Promotion Initiative 2018).

The social and equity categories had fewer and pragmatically more difficult indicators to select among (Appendix D, Table D2), as might be expected from their newer status as additions to the sanitation value chain (Cookey et al. 2022). Studies deemed the aspects of worker safety and access to on-site sanitation services among vulnerable populations as the most globally relevant indicators in line with professionalization and human rights aspects of sanitation work. Worker safety adherence may require internal or external audit procedures to complete rather than being self-reported. Stakeholders should tailor the measurement approach for disaggregating the service access indicator to the diversity of local vulnerable populations (AMCOW 2021; UN Human Rights Council 2017).

Finally, stakeholders cannot ignore finance and management indicators as they are a critical part of the enabling context for functional on-site sanitation services. The finance indicator calls attention to the central role of operating cost recovery in the FSM business model, which has traditionally been more challenging for sanitation than water and more challenging for on-site sanitation than sewered sanitation. The management indicator recognizes the transition to formally regulated services that may better safeguard human health. Like the social and equity indicators, the indicator frames the ongoing global transition to considering on-site sanitation services as a legitimate and permanent sanitation solution (alongside sewered sanitation) where both worker career opportunities and customer satisfaction contribute to the holistic wellbeing of the industry.

Challenges

Among the prioritized indicators, some categories were not represented as priorities, including transport, reuse, gender equity, and climate impacts. Some case study respondents indicated that reporting sludge collection volumes before and after travel to the treatment plant would be redundant, while respondents are unlikely to self-report spills and illegal dumping, especially if a service provider does so intentionally to save on tipping fees. For reuse, the limited scale of activity plus the need for more research on safety standards and regulations forestalled indicator prioritization, although reuse goals may ultimately become more prominent to support a circular economy. Although practiced since 2015, the containment, collection, treatment, disposal, and reuse percentages characterizing on-site sanitation within shit-flow diagrams have often been assessed at a single time point with intensive data collection by a third party, whereas service providers would need to measure the FSM KPIs repeatedly (e.g., once per year or more frequently) to support use in active benchmarking.

Collecting household-level data on containment facilities and disaggregating indicators by population vulnerability (e.g., low-income, women and girls, people with disabilities) may pose a difficulty, especially

for dense urban, peri-urban, or low-income areas where fluctuations in population is common (Carolini and Raman 2021). Remote mapping methods (Atieno, Stuart, and Setty 2023), collaboration with existing household surveys (e.g., census, UNICEF Multiple Indicator Cluster Surveys), postal service address directories (being trialed in Kampala City), or scheduled desludging visits (CWIS 2021a) could potentially simplify maintaining up-to-date household-level inventories. While important to organizational performance, social responsibility KPIs related to gender inclusion in the workforce are challenging to apply consistently across all cultural contexts and labor types. They further risk masking tokenism in practice.

Some KPIs affecting FSM service providers may fall under the jurisdiction of another entity (e.g., an occupational health and safety administration for worker safety, an environmental protection agency for effluent water quality). In these cases, organizations might lack clear information about which party has responsibility for monitoring, leaving the aspect unattended. An example from Narsapur in India noted differences in reporting requirements to the state environmental regulator between public and donor-funded treatment plants (CWIS 2021c). The emerging impetus to integrate climate adaptation into water and sanitation services suggests attention to a KPI such as greenhouse gas emissions from sludge transport trucks or treatment plants, although it would be difficult to estimate widely and accurately, pending expansion of more example efforts and guidance.

Finally, financial sustainability indicators for FSM may prove more difficult to assess relative to water or wastewater as these efforts often came later and piecemeal (e.g., add-ons to existing departments, mandates, budgets, loans, grants, cross-subsidization, partnerships), rather than service providers integrating them within institutional frameworks early on. The Zambia case study suggests that a transition phase and regulatory support can help to institutionalize benchmarking initiatives into routine reporting (CWIS 2021b).

7.1.3 COMPARISON WITH IBNET INDICATORS

Our study prioritized two FSM benchmarking indicators in common with six being piloted by the New IBNET in consultation with CWIS (i.e., workers following standard operating procedures and formal regulation of local pit- or tank-emptying service providers). Although not included in their FSM KPIs, two more (service coverage and operating cost recovery) align with the New IBNET's planned water and sewerage KPIs. Explicit climate adaptation indicators were also pending for the New IBNET as they intended to fold climate action within water conservation (e.g., non-revenue water reduction) and green infrastructure (e.g., cost savings through energy efficiency). Beyond these similarities, we observed the following differences:

Our KPI assessment excluded the highest priority indicator for the New IBNET, having policy, strategy, or regulations for on-site sanitation developed and operationalized, as a situation assessment approach that is unlikely to change year-over-year. Further, the extent of policy implementation would be difficult to assess in a binary (yes/no) fashion. Nevertheless, it is captured as a building block that enables FSM benchmarking (Figure 1). Second-tier indicators considered but not piloted for the New IBNET also largely centered on regulatory and management regimes while we prioritized sludge collection, treatment, and disposal volumes.

Another indicator being piloted by the New IBNET is completion of strategic sanitation planning exercises (yes/no, number). This is most similar to the indicator we assessed regarding strategic use of KPI data in budgeting and investment decisions, which we deemed a lower priority given the varied approaches to evidence use in decision-making and potential subjectivity in assessing decision quality. For example, planned directions may be biased due to stakeholder exclusion, politically motivated, or

agreed but not carried out in practice. Fruitful use of more holistic planning approaches, such as national strategies or city-level sanitation safety plans, might be resource-intensive to assess well (e.g., requiring a thorough annual audit) but could offer a better indication of the influence of planning exercises in practice.

While prioritized for piloting by the New IBNET, this study did not universally recommend having performance agreements in place as they may work differentially in different scenarios and demotivate actors if not well targeted (Howard and White 2020). In contexts not well suited to performance-based agreements, stakeholders could apply other types of incentives. We did, however, recommend incentivizing professionalization through formal licensing and contracts.

Other indicators being piloted by the New IBNET include the average tariff, pit-emptying fee and sludge/septage tipping fee (at the treatment plant). While these might be simple to track, we did not consider them as relevant as cost-recovery to the service providers' financial health (finance indicator), nor as relevant as access (equity indicator) to vulnerable consumers' prospects. Measuring consumer affordability (e.g., tariff as a percentage of annual household expenditures) across wealth quintiles requires more complex categorization of household income or proxy wealth data but could provide a better indication of whether FSM tariffs impinge on the ability of any residents to meet other household needs.

7.1.4 INDICATORS FOR SERVICE PROVIDERS

Service providers should contribute to gathering data on the same KPIs as those used at the city level for consistency (Table 6), although they may opt to track additional FSM KPIs for other purposes (Figure 2). The literature review highlighted a few examples of KPIs that service providers such as smaller utilities or private emptiers might commonly use within their service domains, including fecal sludge collection ratio, safe transport, quality of treated fecal sludge, sludge reuse ratio, customer satisfaction, and worker safety. Table 7 lists several indicators that may be useful to FSM service providers to monitor the quality of their services, financial status, and/or customer satisfaction. An extensive list of possible FSM KPIs appears in Appendix D (Table D2), and their appropriateness and feasibility for use by service providers will vary among contexts.

Table 7: Potential indicators that could be used by FSM service providers

Category	KPI	Definition
Collection	Service area coverage	Percent of population or spatial area that has access to desludging services
	Fecal sludge collection ratio	Volume collected: Estimated volume accumulated
	Worker safety (during desludging)	Provision of protective materials and recommended safety protocols, as opposed to worker adherence
Transport	Rate of safe transport to designated points	Volume delivered to designated points, divided by volume collected
	Frequency of safe transport of fecal sludge	Spillage and/or illegal dumping events per year or per 10,000 trips
	Desludging vehicle maintenance	Percentage of desludging vehicles that comply with maintenance standards; could measure via inspection and maintenance service regularity.
Treatment	Treatment facility utilization	Ratio of volume treated to treatment capacity (in Bangladesh, annual fecal sludge volume treated, divided by design capacity of fecal sludge treatment plant)

Category	KPI	Definition
	Quality of treated fecal sludge	Percentage of tests that meet standards for biosolids, effluent, emissions, and proper waste disposal
	Total volume of sludge treated	Volume of sludge processed by treatment facilities
Disposal/ Reuse	Fecal sludge reuse ratio	Volume of reused sludge (solids) as a percentage of the total volume treated
Customer Satisfaction	Response time to requested desludging	Percentage of requests serviced within 48 hours (India) average time between application and delivered service (Bangladesh)
	Customer satisfaction	A 75% or higher rating on single-question survey (includes pricing satisfaction)
	Customer retention rate	Average percentage of customers using the service year over year
Finance	Average tariff or pit-emptying fee	Average emptying fee charged to household customers
	Customer conversion	Acceptance rate of offers for emptying
	Bill collection efficiency	Percentage of revenue received divided by revenue owed for services
	Staff efficiency	Average personnel cost per number of staff per month

7.1.5 TARGET SETTING

Target goals often accompany KPIs, although values can still be compared to one's performance over time or others' performance without setting targets. If used, moving incremental performance targets should be tailored to the local system. From the case studies, the experience with Central Finance Commission targets in India recognized that static targets were not universally feasible, sparking increased customization. The metropolitan regulator in the Philippines noted the relative ease and acceptability of performing well on smaller tasks. For example, they did not expect to observe failures on treatment indicator targets while infrastructure remains in near-new condition. In Haiti, the service provider considered achievement of one global Sustainable Development Goal for sanitation access demoralizing. However, they could still demonstrate progress using customized targets for the local context.

7.2 GENDER EQUITY AND SOCIAL INCLUSION INDICATORS

Benchmarking risks taking a narrow view of performance that over-emphasizes operational and financial indicators and neglects social and environmental concerns. Although not necessarily intended, interpretation of benchmarking may equate commercial metrics with overall performance (McDonald 2016). Incorporating performance indicators dedicated to social aspects can help mitigate this risk (M. Mehta, Mehta, and Immanuel 2011; WSUP 2015). Several emerging KPIs relate to gender equity and social inclusion in the water and wastewater sector (Appendix D, Table D1) and for FSM (Appendix D, Table D2), although these remain absent from most benchmarking initiatives. For instance, the World Bank's IBNET initiative tracks the proportion of female workers in water utilities, as well as women's average salaries globally (World Bank 2019). Additional research and practitioner experience will help to identify and refine the most appropriate indicators.

Some utilities have also trialed or adopted additional KPIs related to gender equity in the workforce. In Ethiopia, Argentina, and Malawi, a global benchmarking report related to gender equity in the workforce led to increased representation of women in decision-making positions for water utilities (Jha 2022; Muximpua and Hatzfeldt 2020; World Bank 2019). The Economic Dividends for Gender Equality

Certification Foundation, which provides an international benchmark on gender equity for companies, assessed opportunities for female employees at water utilities in Albania, Kosovo, and Romania (World Bank 2019). So far, the World Bank has driven these expanded gender benchmarking initiatives, which are not yet incorporated into domestic monitoring and regulatory systems.

With respect to social inclusion, countries such as Kenya, India, and Uganda have introduced indicators for water and wastewater services focused on serving the poor, while the CWIS Initiative and others have proposed similar pro-poor FSM indicators (Appendix D, Table D2). In 2018, the Kenyan regulator WASREB introduced "KPI 10," defined as the percentage of service coverage in low-income areas. Before this, a two-year pilot in the largest nine utilities helped confirm the feasibility and benefits of tracking this new KPI. However, utilities struggle to maintain up-to-date information on the location and borders of fast-evolving low-income areas, making accurate measurement of KPI 10 difficult (WSUP 2018). Similarly, in India, a few states have a separate indicator quantifying access to services in slums (M. Mehta, Mehta, and Immanuel 2011). In Uganda, the NWSC has held a performance contract with the state since 2000, which requires the utility to extend services to the poor. This requirement appears in the utility's internal performance contracts with area managers, which set explicit and measurable targets to ensure that the poor receive services. Linked to these targets, the NWSC has established several pro-poor measures, including network expansions to low-income areas, densification of water kiosks, reduced connection fees for poor customers, and a 10 percent surcharge on tariffs to fund this subsidy (Sekayizzi, Odonga, and Amayo 2005).

7.3 COMPOSITE INDICATORS

Benchmarking sometimes uses composite indicators to summarize multiple KPIs into a single score, simplifying rankings or comparisons among service providers (Strande et al. 2018; ESAWAS 2019). For example, the IBNET "Apgar" score for water and sanitation utilities calculates a composite score from five KPIs (or six, if the utility also provides sewerage services). It summarizes information capturing the utility's operational, financial, and social performance. Composite scores mainly offer the benefit of easing communication to the public and decision-makers, when compared to many separate indicators (a dashboard approach). Their main drawback is that the respective weights or values placed on individual indicators are subjective and prone to political biases (S. Berg 2020; Gallego-Ayala et al. 2014). Composite indicators can also mask large deficiencies in individual indicators.

To address these shortcomings, "spiderweb" evaluation frameworks, as shown in Figure 3, have been proposed as an alternative to composite indicators. This type of display clearly communicates performance levels relative to any specific component (length of each edge), as well as overall performance (the shaded area).

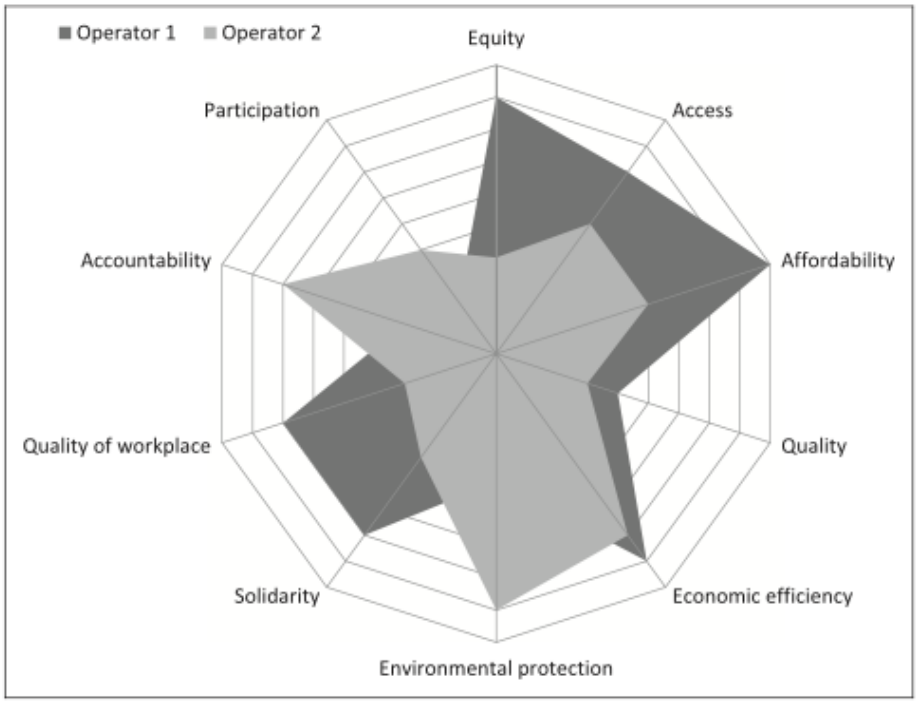

Figure 3: Visual "spiderweb" approach to compare service providers (McDonald 2016)

When comparing KPIs among service providers, the service area size substantially affects several measures (e.g., non-revenue water, percent service coverage, low-income coverage, operating costs). As a result, direct comparison of small and large service providers may not be meaningful. To address this, the Kenyan WASREB categorizes service providers by size (small, medium, large, and very large) and provides rankings by category in addition to general rankings (WASREB 2022). The IBNET also groups utilities by their service population size class (i.e., <10,000; 10–50,000; 50–100,000;100–500,000; 500,000–1 million [M], >1M) (Danilenko et al. 2014).

8.0 BENCHMARKING OUTCOMES

While benchmarking serves as a critical tool for monitoring and improving the performance of service providers (e.g., S. Berg 2010; 2020; WSUP 2018; Sinha 2013; Minelli 2021), there have been limited studies evaluating the impact of benchmarking on service delivery performance (McDonald 2016; Goh and See 2020). Literature from the healthcare sector similarly flagged the lack of evidence on benchmarking effectiveness (Wait and Nolte 2005). Benchmarking requires a certain level of institutional capacity, staff training, and robust data systems to measure KPIs. As such, benchmarking datasets may over-represent well-resourced service providers and exclude lower-capacity service providers, regardless of their actual service performance.

8.1 OUTCOME ASSESSMENT CHALLENGES

Benchmarking does not easily lend itself to controlled prospective experiments, which would allow for attribution of impacts. The typically large-scale units of intervention (city or national service providers) make it difficult to identify comparable controls. Analysis of natural experiments could provide helpful information, but this review identified none. Differences among service providers' measurement approaches and data tracking variations within the same service provider over time can confound comparative analyses (Ganjidoost et al. 2018). Some evidence shows discrete programmatic decisions stemming from global FSM benchmarking tools such as shit-flow diagrams, although most data uses have been strictly academic (Furlong, Mugendi, and Brdjanovic 2023). Before-after comparisons would be difficult, because by definition, KPI data is generally scarce before the introduction of benchmarking.

Clearly, disentangling the effect of benchmarking itself from other improvement efforts taking place simultaneously (e.g., infrastructure upgrades, staff changes, new funding) within the same water or sanitation system poses a challenge. Retrospective observational studies have a limited ability to attribute changes specifically to the benchmarking program as opposed to any other factors, but evidence showing that data triggered certain management responses and subsequent changes is certainly supportive. Some longitudinal studies find improvements in performance following the onset of KPI monitoring. A study in France, for example, reported a reduction in customer complaints and increased water quality sensor availability after the utility started tracking these KPIs for operational purposes (Setty et al. 2018). Longitudinal analysis of available benchmarking data does not universally reveal improvements in performance over time, though, as corresponding management improvements may not be implemented, short-term initiatives may not be sustained, a performance "ceiling" may be reached, or other contextual factors may influence outcomes. Example data on a non-revenue water KPI from four Kenyan water service providers in 2008–2022 reveal no unifying trend over time (Figure 4).

URBAN WASH previously found that cities that had historically strong progress toward water and sanitation access used performance indicator tracking. However, the analysis could not ascertain the extent to which the act of benchmarking versus other factors such as champions led directly to improvements and success (URBAN WASH 2023). Service providers could apply another qualitative method called "process tracing" to rigorously assess the link between an outcome of interest and an explanation, weighting of evidence for and against causal inference (Ricks and Liu 2018), but this method only allows them to trace a small number of in-depth cases given the time and resources needed to gather sufficient detail (Beach 2017).

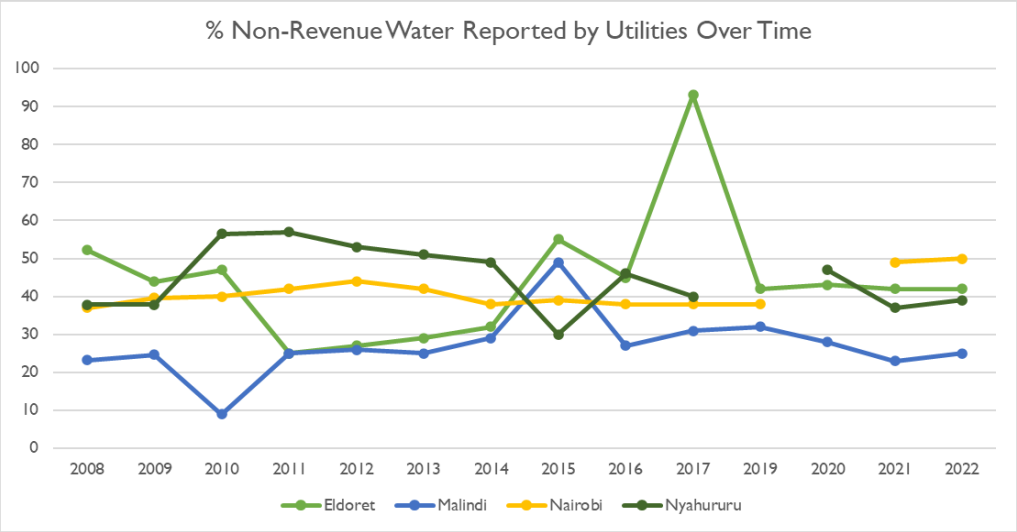

Figure 4: Example results of a 2008–2022 KPI from four Kenyan water service providers

8.2 COMMON BENEFITS

Despite the limited scientific evidence, a number of common-sense arguments explain the general consensus that water and wastewater benchmarking improves performance. The first argument asserts that "you can't manage what you don't measure" (e.g., Vavaliya and Bhavsar 2022). Performance indicators, when used appropriately as part of the decision-making process, can inform effective management. Relatedly, transparent KPI data promote accountability, a pillar of good governance (GIZ 2019; URBAN WASH 2023). Therefore, benchmarking can offer a fundamental first step toward strengthening service provider capacities, when combined with effective oversight, data sharing, and continuous improvement efforts.

Second, service providers often deploy benchmarking programs in combination with "carrots and sticks" known to influence institutional performance: incentives (financial and reputational) to reward good performance and/or penalties in response to poor performance (e.g., fine imposed by the regulator or non-payment in the context of a public-private partnership). Of note, performance contracts linking benchmarking data with explicit penalties or incentives played a substantial role in shifting the performance of both the National Office of Water and Sanitation in Burkina Faso and the NWSC in Uganda, two top-performing African utilities (GIZ 2019).

Third, peer-to-peer comparisons appeal to feelings of pride and friendly competition among utility leaders and can cause underperformers to emulate good practices. One Canadian report explained, "while the initial intent of the project was metric benchmarking, the project has evolved into a dynamic vehicle for the development, sharing, and implementation of municipal best practices" (Atherton 2023). Similarly, in the Netherlands, efficiency and customer satisfaction increased following the introduction of voluntary benchmarking (EurEau 2015).

8.3 COMMON CRITIQUES

Despite the dominant view that benchmarking contributes to the enabling environment for effective services, critiques have highlighted a few concerns. Notably, compiling and submitting good-quality performance data for benchmarking requires considerable capacity and resources. Monitoring requirements may further marginalize under-capacitated service providers in low-resource settings

(McDonald 2016). The healthcare literature also warns against overreliance on KPI data, which often draws from estimates rather than real measurements (Wait and Nolte 2005). The electricity literature similarly flags data errors as a primary limitation of benchmarking (Jamasb and Pollitt 2000).

The focus on pre-defined indicators may lead to limited views of performance and take bandwidth away from creative problem-solving. "At its worst, instead of being a force for change, benchmarking can put a stop to serious analysis of problems and/or experimentation with [locally defined] innovative solutions" (Sisson, Arrowsmith, and Marginson 2003). A similar critique emerged from the healthcare literature, in that benchmarking indicators may focus on available data rather than improvement priorities, taking attention away from problem-solving (Wait and Nolte 2005). Additionally, benchmarking may impose Western-centric views of performance (Sisson, Arrowsmith, and Marginson 2003), and the emphasis that international donors put on benchmarking may reflect the persistence of colonial attitudes (McDonald 2016). Further, when implemented in a top-down manner, benchmarking does not give voice to lower-level employees or consumers and risks neglecting social and environmental concerns (McDonald 2016).

8.4 OUTCOMES OF FECAL SLUDGE MANAGEMENT BENCHMARKING

The case study findings provided novel evidence regarding the ways benchmarking can assist or detract from service goal achievement. A cyclical relationship appeared wherein FSM benchmarking data can drive programming and FSM programming can advance data collection. As with the "chicken and egg" problem, each may trigger or perpetuate the other. In Kenya, for example, the WHO/UNICEF JMP and Kenya's Vision 2030 drove recognition of a lack of progress toward sanitation targets via sewerage and the need for a national FSM policy.

The FSM case studies demonstrated that benchmarking clearly represents a versatile tool with many uses, although its mode of implementation must carefully align with its intended goals to avoid unintended consequences. Negative experiences with international donor initiatives and data-sharing collaborations likely go underreported (Evertsz, Bull, and Pratt 2023). During the interviews, most respondents initially reported only positive outcomes of benchmarking. After probing, however, they went on to observe a relatively equivalent range of positive and negative outcomes of benchmarking activities (Table 7; Table 8). The range of reported outcomes highlights the importance of considering the potential downsides and opportunity costs of introducing novel, albeit promising, interventions that might have functioned well in other contexts.

While improving program performance represented an expected beneficial outcome stemming from the definition of benchmarking, a more novel finding concerned the prevalent use of benchmarking to attract financing (Table 8). Respondents indicated that having benchmarking data to report often satisfied past funders and made their organization appear more trustworthy and creditworthy to future funders and creditors. From the donor perspective, financial and performance KPI data may be used to determine eligibility for borrowing or to segment markets for optimal resource allocation (WASH-FIN 2021; Peal et al. 2015). This benefit was more well understood among respondents, compared to documentation in the literature. Other unanticipated benefits included the potential use of benchmarking data for scenario forecasting (e.g., of different potential future management approaches) and enhancing support and buy-in for FSM activities among managers and board members.

Table 8. Range of positive self-reported outcomes from FSM benchmarking case studies*

Outcome	Glob	Ug	Za	Ba	Indi	Ph	Ha
Drives programming performance improvement	+	+	+	+	+		
Attracts funding or financing	+	+		+			+

Outcome	Glob	Ug	Za	Ba	Indi	Ph	Ha
Aids tailoring programs to context	+	+	+				
Heightens accountability			+	+		+	
Tracks incremental change to set viable goals			+		+	+	
Communicates impact (within organization and publicly)		+	+				
Motivates competition			+			+	
Supports learning and decision-making			+				+
Improves data reporting culture					+		+
Facilitates forecasting investment outcomes	+						
Enhances management and board support			+				

*Glob = Global; Ug = Uganda; Za = Zambia; Ba = Bangladesh; Indi = India; Ph = Philippines; Ha = Haiti; Few outcomes reported to date for Kenya and Indonesia

Regarding negative outcomes (Table 9), respondents reported fewer issues with operators refusing to cooperate with KPI data collection, compared to the literature. Instead, they highlighted a range of potential challenges, including the opportunity cost, inability to capture a complete picture using only quantitative indicators, and the potential interpersonal aspects of working for an organization that might overly value benchmarking or weigh indicators in a way that does not align with employees' personal values. The theme of the organization's control over performance outcomes also arose, with respect to other outside factors that have potential to influence the data to a greater degree but may not be easy to document. In addition, respondents described a need to set targets correctly to avoid complacency (e.g., lack of effort from targets set too low) or, conversely, motivate false reports (e.g., from targets set unachievably high).

Table 9. Range of negative self-reported outcomes from FSM benchmarking case studies*

Outcome	Glob	Ug	Za	Ba	Indi	Ph	Ha
Distracts from other issues and needs		+	+		+	+	
Requires skills, capacity, and time	+	+		+			
Does not cover all sanitation aspects			+		+	+	
Invades professional privacy		+					+
Encounter negative perceptions differing by context and personal interactions (e.g., with benchmarking terms, data uses)	+						
May lead to complacency (e.g., targets require little effort to meet, no new goal areas pursued)		+					
May distribute labor unequally (e.g., not all actors contribute the same effort to benchmarking activities)		+					
Raises concerns about interpreting how externalities influence the data (e.g., political or environmental events)			+				
Contributes to excessive or inconsistent external data requests (e.g., paper and digital, different indicator sets)			+				
Drives cheating (e.g., data appear rosier than conditions observed by residents)			+				
Affects staff and customer morale (e.g., from prioritizing financial performance over equity)							+

**Glob = Global; Ug = Uganda; Za = Zambia; Ba = Bangladesh; Indi = India; Ph = Philippines; Ha = Haiti; Few outcomes reported to date for Kenya and Indonesia.

9.0 RECOMMENDATIONS

Benchmarking has contributed to fostering performance improvements and strengthening the enabling environment for effective water and wastewater services and is now expanding to FSM. Through a literature review and case studies, this review sought to understand good practices for FSM benchmarking and provide recommendations to those seeking to implement FSM benchmarking in different institutional and governance contexts.

BENEFITS OF FSM BENCHMARKING

Benchmarking can support advancements in on-site sanitation service provision but does not represent a silver bullet. FSM benchmarking efforts reported a wide range of both positive and negative outcomes beyond performance improvement. Thus, development partners should not assume benchmarking represents the next or only initiative needed to improve FSM service performance. Further, service providers must carefully measure the degree of benchmarking activity to maximize potential benefits, given that a high number of KPIs will likely increase the opportunity cost of benchmarking. Where practitioners lack resources to make improvements, they might streamline monitoring efforts to focus on financial sustainability. Benchmarking has the potential to improve efficiency, working conditions, and service delivery, optimizing spending in the long run.

ADVANCING ADOPTION AND USE OF BENCHMARKING

Successful adoption of FSM benchmarking assumes existing basic elements of on-site sanitation, organizational and contextual readiness, and clear communication channels. Once these conditions are in place, potential entry points for undertaking FSM benchmarking include regulatory ownership, customer demand, capacity strengthening via data champions, and financial support. The collective learning curve to ensure support for all aspects of benchmarking will differ among contexts. In challenging contexts (e.g., those lacking regulatory support), practitioners can reduce effort on benchmarking to those components that offer value while advocating to address other critical program development aspects.

To establish data champions, implementers can assign key roles and responsibilities for each actor involved in benchmarking and convene regular meetings among a network working toward similar goals. Those championing benchmarking activities can help to match solutions to implementation barriers as they arise (Waltz et al. 2019). Where data access is restricted, responsible parties can clarify the underlying reasons and advocate for leaders to issue guidance and incentives that promote wider data sharing. Logistically, it may be helpful to clarify which specific management questions the data is used to address and assign trained staff to carry out data visualization in an agreed reporting format built around stakeholders' decision needs.

Effective data management, incentives, and public data dissemination foster beneficial outcomes of benchmarking activities. These good practices interrelate, such that incentives and data management contribute to the effectiveness of public data sharing and vice-versa. A short list of shared indicators, behavior reinforcement techniques (e.g., rewards, penalties, licenses, contracts tailored to local needs), and data management systems (including automated support) help to maintain effective benchmarking systems. To avoid gathering unusable data, practitioners should strategize about how to transparently track data revisions and integrate simplified validation or quality assurance measures. Layered benchmarking systems (e.g., to support multiple programs) must take care to avoid redundant or incompatible data requests.

Combining multiple forms of benchmarking, such as using process or customer survey benchmarking with metric benchmarking, can create synergies. Pioneer or "model" systems help to support process benchmarking, replication, and scale-up. Where no demonstration projects exist, practitioners might identify a similar location in another region and conduct peer learning site visits before starting benchmarking initiatives. For example, Kampala and Lusaka cities matched with Durban for benchmarking visits under the Reinforcing African Sanitation Operators Partnership (Senkwe and Niwagaba 2016). Interviewees proudly reported hosting peer-learning visits from other cities' utilities, and one respondent shared a preference to "learn by seeing." If the highest priority performance issues remain unclear from metric KPI data alone, broad stakeholder input from employees and customers can provide another line of evidence to help interpret it.

FSM BENCHMARKING METRICS

Benchmarking activities should engage global, regional, and national networks, where possible and appropriate, to maintain comparability and offer opportunities for knowledge exchange and learning. Critically, feedback loops among parties providing and using benchmarking data, as well as opportunities for public input, benefit proper data reporting and use. Global and regional professional networks (e.g., SuSanA, FSM Alliance, IWA, ESAWAS, PWWA, and/or the new African Union Regional Economic Communities) and national associations and networks can play a helpful role in standardizing and disseminating recommendations. These communication pathways strengthen the effectiveness of benchmarking activities by drawing meaning from the data and facilitating appropriate response actions.

For FSM benchmarking, KPIs should encapsulate the full sanitation value chain, including social, financial, and environmental responsibility. Over time, sticking to consistent, reusable indicators that cover the full sanitation service chain and externalities will aid peer comparison as well as long-term trend evaluation. Benchmarking systems can apply at many scales and typically use fewer indicators or aggregate indicators across larger spatial scales. We propose nine indicators for comparison among cities, stemming from their existing use, benefits, available alternatives, measurement challenges, risk of bias, and other factors: 1) percentage of population with access to safely managed sanitation, 2) access among vulnerable groups, 3) coverage of on-site sanitation facilities, (4) total sludge collection volume, 5) treatment ratio, 6) safely managed disposal or reuse, 7) worker safety procedure adherence, 8) operating cost recovery, and 9) percentage of desludging providers operating with a formal license.

Service providers and implementers should customize FSM KPIs for their location and contextual needs, revisiting the benchmarking approach and indicators periodically. This may involve adjusting KPIs and targets in consultation with local stakeholders familiar with the needs on the ground. Service providers at the local level might add other indicators, for example addressing desludging methods, customer acquisition, or worker efficiency. Indicator selection that aligns with both a) external data sharing purposes and b) internal organizational needs can help to address various actors' motivations for benchmarking while leveraging limited resources to support potential secondary data uses.

10.0 REFERENCES

Accenture. 2013. "Reflections on Performance Benchmarking in the Dutch Drinking Water Industry." REFLECTIONS ON PERFORMANCE 2012. Association of Dutch Water Companies.

ADB. 2022. "A Governance Approach to Urban Water Public–Private Partnerships: Case Studies and Lessons from Asia and the Pacific." Manila, Philippines: Asian Development Bank. https://doi.org/10.22617/SPR220100.

AECOM. 2012. "National Solid Waste Benchmarking Initiative." Canada. https://nswbi.nationalbenchmarking.ca/docs/reports/NSWBI_2012_Report-Public_version.pdf.

African Development Bank, United Nations Environment Programme, and GRID-Arendal. 2020. "Sanitation and Wastewater Atlas of Africa." Abidjan, Nairobi and Arendal: AfDB; UNEP; GRID-Arendal. https://www.unep.org/resources/publication/sanitation-and-wastewater-atlas-africa.

Al-Muyeed, Abdullah, Sanjoy Mukherjee, and Md Shafiqul Hassan. n.d. "Scaling up CWIS towards Achieving SDG 6.2 in Bangladesh." International Water Association. Accessed August 7, 2023. https://iwa-network.org/iusstory-bangladesh/.

AMCOW. 2021. "African Sanitation Policy Guidelines." Abuja, Nigeria: African Ministers' Council on Water. https://amcow-online.org/initiatives/african-sanitation-policy-guidelines-aspg.

American Water Works Association. n.d. "AWWA History." Accessed February 23, 2024. https://www.awwa.org/About-Us/AWWA-History.

Aspire to Innovate (a2i) Programme. 2023. "Sanitation Data Command Center: The First of Its Kind in South Asia." 2023. https://a2i.gov.bd/sanitation-data-command-center-the-first-of-its-kind-in-south-asia/.

Athena Infonomics. 2020. "Monitoring Tools for Fecal Sludge Collection & Transport." CWIS Learning Brief Series. CWIS.

Atherton, Jeff. 2023. "The Canadian Infrastructure Benchmarking Initiative (CIBI): Setting a Global Standard." *digitalAECOM* (blog). 2023. https://digital.aecom.com/project/canadian_infrastructure_benchmarking_initiative.

Atieno, Irene, Kara Stuart, and Karen Setty. 2023. "Mapping the Unmapped: Why Understanding Low-Income Areas Is Essential." *Aquaya* (blog). March 20, 2023. https://aquaya.org/mapping-the-unmapped-why-understanding-low-income-areas-is-essential/.

AWWA. 2019. "Key Performance Indicators for Non-Revenue Water." American Water Works Association. https://www.awwa.org/Portals/0/AWWA/ETS/Resources/WLCCKPIReport%202019.pdf?ver=2019-11-20-094638-933.

——— 2020. *AWWA Utility Benchmarking: Performance Management for Water and Wastewater*. Benchmarking Data from 2019 for 58 Key Performance Indicators. United States of America.: AWWA.

Bartram, Jamie, and Karen Setty. 2021. "Environmental Health Science and Engineering for Policy, Programming, and Practice." *Journal of Environmental Engineering* 147 (10): 03121002. https://doi.org/10.1061/(ASCE)EE.1943-7870.0001880.

Basyal, Isha. 2023. "Citywide Inclusive Sanitation in Odisha: A Journey of Empowerment and Transformation." *Global Water and Sanitation Center* (blog). 2023. https://www.gwsc.ait.ac.th/citywide-inclusive-sanitation-in-odisha-a-journey-of-empowerment-and-transformation/.

Beach, Derek. 2017. "Process Tracing Methods in the Social Sciences." In *Oxford Research Encyclopedia of Politics*, by Derek Beach. Oxford University Press. https://doi.org/10.1093/acrefore/9780190228637.013.176.

Berendes, David M., Trent A. Sumner, and Joe M. Brown. 2017. "Safely Managed Sanitation for All Means Fecal Sludge Management for At Least 1.8 Billion People in Low and Middle Income Countries." *Environmental Science & Technology* 51 (5): 3074–83. https://doi.org/10.1021/acs.est.6b06019.

Berg, Caroline van den, and Alexander Danilenko. 2011. *The IBNET Water Supply and Sanitation Performance Blue Book*. World Bank. https://doi.org/10.1596/978-0-8213-8582-1.

——— 2017. *Performance of Water Utilities in Africa*. Washington, DC: World Bank. https://doi.org/10.1596/26186.

Berg, Sanford. 2010. *Water Utility Benchmarking: Measurement, Methodologies, and Performance Incentives*. London: IWA Publishing.

——— 2013. "Best Practices in Regulating State-Owned and Municipal Water Utilities." Project document. Santiago, Chile: Economic Commission for Latin America and the Caribbean (ECLAC).

——— 2020. "Performance Assessment Using Key Performance Indicators (KPIs) for Water Utilities: A Primer." *Water Economics and Policy* 06 (02): 2050001. https://doi.org/10.1142/S2382624X20500010.

Berg, Sanford, and Julie C Padowski. 2010. "Overview of Water Utility Benchmarking Methodologies: From Indicators to Incentives." USA: Public Utility Research, University of Florida. https://bear.warrington.ufl.edu/centers/purc/docs/papers/0712_Berg_Overview_of_Water.pdf.

Blankenship, Linda, Myron Olstein, and Barry Liner. 1998. "Metric Benchmarking." *American Water Works Association* 90 (6).

Bulson, Paul, Susilawati Muslimah, Nicolas Reynaud, Isabel Blackett, Alexander Viwat Campbell, and Inni Arsyini. 2021. "Independent Evaluation of SANIMAS Model as an Approach for Providing Decentralised Sanitation." Citywide Inclusive Sanitation Technical Assistance Hub for South Asia. https://www.isdb.org/sites/default/files/media/documents/2021-03/SANIMAS%20Model_as%20an%20Approach%20for%20Prodiving%20Decentralised%20Sanitation_March%202021.pdf.

Cabrera, Enrique, Peter Dane, Scott Haskins, and Heimo Theuretzbacher-Fritz. 2011. *Benchmarking Water Services: Guiding Water Utilities to Excellence*. https://doi.org/10.2166/9781780400877.

Carolini, Gabriella Y., and Prassanna Raman. 2021. "Why Detailing Spatial Equity Matters in Water and Sanitation Evaluations." *Journal of the American Planning Association* 87 (1): 101–7. https://doi.org/10.1080/01944363.2020.1788416.

CEPT University. 2010. "Performance Measurement Framework for Urban Water and Sanitation. Volume 1: Approach and Framework."

——— 2015. "SAN Benchmarks: Citywide Assessment of Sanitation Service Delivery Including on-Site Sanitation." https://pas.org.in/Portal/document/UrbanSanitation/uploads/SanBenchmark.pdf.

——— 2018. "Performance Assessment System for Urban Water Supply and Sanitation." 2008–2016. Ahmedabad, India: CEPT University: Center for Water and Sanitation (C-WAS): Performance Assessment System (PAS).

Cookey, Peter Emmanuel, Thammarat Koottatep, Walter Thomas Gibson, and Chongrak Polprasert. 2022. *Integrated Functional Sanitation Value Chain: The Role of the Sanitation Economy*. 3rd ed. London, UK: IWA Publishing. https://www.iwapublishing.com/books/9781789061833/integrated-functional-sanitation-value-chain-role-sanitation-economy.

Corton, Maria Luisa, and Alejo Molinari. 2007. "Regulatory Collaboration in the Americas: ADERASA." https://bear.warrington.ufl.edu/centers/purc/docs//papers/0730_Corton_Regulatory_Collaboration_in.pdf.

CWAS. 2021. "Performance Assessment System: Interactive Dashboards." 2021. https://pas.org.in/web/ceptpas/interactivedashboards?p_p_id=InteractiveDashboard_WAR_Portal&p_p_lifecycle=1&p_p_state=normal&p_p_mode=view&p_p_col_id=column-1&p_p_col_count=1&actionVal=Retrieve&SkipAccessChecking=false.

CWIS. 2021a. "Sanitation Data Ecosystem Profile: Kampala City." https://s3.amazonaws.com/resources.cwis.com/learning/234/KampalaCitySanitationDataEcosystemReport.pdf.

——— 2021b. "Sanitation Data Ecosystem Profile: Lusaka City." https://s3.amazonaws.com/resources.cwis.com/learning/240/LusakaCitySanitationDataEcosystemReport.pdf.

——— 2021c. "Sanitation Data Ecosystem Profile: Narsapur City." https://s3.amazonaws.com/resources.cwis.com/learning/237/NarsapurCitySanitationDataEcosystemReport.pdf.

"CWIS Measurement Note." n.d. Accessed November 13, 2023. https://www.cwiscities.com/Dashboard/DashboardInfo.

Danilenko, Alexander, Caroline van den Berg, Berta Macheve, and L. Joe Moffitt. 2014. *The IBNET Water Supply and Sanitation Blue Book 2014: The International Benchmarking Network for Water and Sanitation Utilities Databook*. World Bank Publications.

De Goede, Marieke, Bert Enserink, Ignaz Worm, and Jan Peter Van Der Hoek. 2016. "Drivers for Performance Improvement Originating from the Dutch Drinking Water Benchmark." *Water Policy* 18 (5): 1247–66. https://doi.org/10.2166/wp.2016.125.

DINEPA, Zami Lasante, and DAI. 2022. "Lessons Learned from 12 Years of Urban Sanitation Work in Haiti." https://www.globalwaters.org/sites/default/files/lessons_learned_in_urban_sanitation.pdf.

Dubois, Alexandra. 2023. "Unlocking the Potential of SFD: Exploring Diverse Applications in GIZ Projects."

Eales, Katherine, and Masroor Ahmad. 2010. "A Review in Bangladesh, India, and Pakistan: Benchmarking for Performance Improvement in Urban Utilities." The Water and Sanitation Program, World Bank. https://www.wsp.org/sites/wsp/files/publications/WSP_benchmarking.pdf.

EBC. n.d. "The European Benchmarking Co-Operation: Your Portal to European Best Practices." European Benchmarking Co-operation. Accessed May 18, 2023. https://www.waterbenchmark.org/style/custom/ebc/download/EBC-Brochure.pdf.

ESAWAS. 2019. "Regional Benchmarking of Water Supply and Sanitation Utilities." 2018/2019 Report.

——— 2021. "Regional Benchmarking of Water Supply and Sanitation Utilities 2020/2021 Report." Eastern and Southern Africa Water and Sanitation Regulators Association. https://www.esawas.org/index.php/publications/regional-benchmarking-of-large-wss-utilities-report?task=download.send&id=73&catid=7&m=0.

——— 2022a. "Regional Benchmarking of Water Supply and Sanitation Utilities: 2021/2022 Report." Eastern and Southern Africa Water and Sanitation Regulators Association. https://esawas.org/index.php/publications/regional-benchmarking-of-large-wss-utilities-report?task=download.send&id=79&catid=7&m=0.

——— 2022b. "The Water Supply and Sanitation Regulatory Landscape Across Africa; Content Wide Synthesis Report." Synthesis Report. Eastern and Southern Africa Water and Sanitation Regulators Association. https://esawas.org/repository/Esawas_Report_2022.pdf.

EurEau. 2015. "How Benchmarking Is Used in the Water Sector." Belgium.

Evertsz, Natalia, Susan Bull, and Bridget Pratt. 2023. "What Constitutes Equitable Data Sharing in Global Health Research? A Scoping Review of the Literature on Low-Income and Middle-Income Country Stakeholders' Perspectives." *BMJ Global Health* 8 (3): e010157. https://doi.org/10.1136/bmjgh-2022-010157.

EWURA. 2020. "Water and Wastewater Quality Monitoring Guidelines for Water Supply and Sanitation Authorities." Second Edition. Energy and Water Utilities Regulatory Authority.

Ferrari, Rosella. 2015. "Writing Narrative Style Literature Reviews." *Medical Writing* 24 (4): 230–35. https://doi.org/10.1179/2047480615Z.000000000329.

Franceys, Richard, Sam Drabble, and Rosie Renouf. 2020. "Referee! Responsibilities, Regulations and Regulating for Urban Sanitation." Discussion Paper. Eastern and Southern Africa Water and Sanitation Regulators Association (ESAWAS) and Water & Sanitation for the Urban Poor (WSUP).

"FS & SWM Dashboard." n.d. Accessed February 23, 2024. http://sanboard.gov.bd/cwis.

Furlong, C., A. Mugendi, and D. Brdjanovic. 2023. "Exploring the Usage and Impact of Urban Sanitation Tools Targeting Low- and Middle-Income Countries." *Frontiers in Environmental Science* 11. https://www.frontiersin.org/articles/10.3389/fenvs.2023.1109306.

Gallego-Ayala, J, C Dimene, A Munhequete, and R Amos. 2014. "Assessing the Performance of Urban Water Utilities in Mozambique Using a Water Utility Performance Index." *Water SA* 40 (4): 665. https://doi.org/10.4314/wsa.v40i4.12.

Ganjidoost, A., M. A. Knight, A. J. A. Unger, and C. T. Haas. 2018. "Benchmark Performance Indicators for Utility Water and Wastewater Pipelines Infrastructure." *Journal of Water Resources Planning and Management* 144 (3): 04018003. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000890.

GIZ. 2019. "Access to Water and Sanitation in Sub-Saharan Africa: Review of Sector Reforms and Investments, Key Findings to Inform Future Support to Sector Development." Synthesis Report.

GLAAS. n.d. "Data." UN-Water Global Analysis and Assessment of Sanitation and Drinking Water, Data Portal. Accessed January 16, 2023. https://glaas.who.int/glaas/data.

Goh, Kim Huat, and Kok Fong See. 2020. "Twenty Years of Water Utility Benchmarking: A Bibliometric Analysis of Emerging Interest in Water Research and Collaboration." *Journal of Cleaner Production* 284: 124711. https://doi.org/10.1016/j.jclepro.2020.124711.

Howard, Guy, and Zach White. 2020. "Does Payment by Results Work? Lessons from a Multi-Country Wash Programme." *Journal of Water, Sanitation and Hygiene for Development* 10 (4): 716–23. https://doi.org/10.2166/washdev.2020.039.

IBNET. 2021a. "Other Benchmarking Methodologies." 2021. https://www.ib-net.org/benchmarking-methodologies/other-benchmarking-methodologies/.

——— 2021b. "Verify Data." IBNET. 2021. https://www.ib-net.org/benchmarking-methodologies/checklist/verify-data/.

Inclusive Sanitation in Practice. n.d. "City Service Delivery Assessment for Citywide Inclusive Sanitation (CSDA)." Inclusive Sanitation in Practice. Accessed June 20, 2023. https://incsanprac.com/tools.html#CSDA.

IWA. 2021. "AquaRating." International Water Association. 2021. https://iwa-network.org/projects/aquarating/.

Jamasb, T, and M Pollitt. 2000. "Benchmarking and Regulation: International Electricity Experience." *Utilities Policy* 9 (3): 107–30. https://doi.org/10.1016/S0957-1787(01)00010-8.

Jayathilake, N., P. Drechsel, B. Keraita, S. Fernando, and M. A. Hanjra. 2019. "Guidelines and Regulations for Fecal Sludge Management from On-Site Sanitation Facilities." Resource Recovery & Reuse Series Issue 14. Colombo, Sri Lanka: International Water Management Institute (IWMI): CGIAR Research Program on Water, Land and Ecosystems (WLE). https://doi.org/10.5337/2019.211.

Jha, Saroj Kumar. 2022. "Climate Adaptation in Action: Global Water Security and Sanitation Partnership." *World Bank Blogs* (blog). November 11, 2022. https://blogs.worldbank.org/water/climate-adaptation-action-global-water-security-and-sanitation-partnership.

JICA. 2022. "Data Collection Survey on Urban Sanitation, Sewerage and Faecal Sludge Management in Africa." Final Report.

KCCA. 2019. "Kampala Fecal Sludge Management: Improving Feacal Sludge Management for on-Site Sanitation." Kampala Capital City Authority. https://www.kcca.go.ug/uDocs/Improving%20feacal%20sludge%20management%20for%20on-site%20sanitation.pdf.

Kenya Ministry of Water, Sanitation and Irrigation. 2021. "Draft National Sanitation Management Policy: Towards Safely Managed Sanitation for All by 2030 and Beyond."

Kenya Ministry of Water, Sanitation and Irrigation and African Population and Health Research Center. 2022. "Development of the National Sanitation Management Policy (NSMP) in Kenya: A Narrative Descriptive Account of the Journey." 2022. https://aphrc.org/wp-content/uploads/2022/12/The-National-Sanitation-Management-Policy-NSMP-Formulation-Journey.pdf.Knee, Jackie, Trent Sumner, Zaida Adriano, Claire Anderson, Farran Bush, Drew Capone, Veronica Casmo, David Holcomb, Pete Kolsky, Amy MacDougall, Evgeniya Molotkova, Judite Monteiro Braga, Celina Russo, Wolf Peter Schmidt, Jill Stewart, Winnie Zambrana, Valentina Zuin, Rassul Nalá, Oliver Cumming, Joe Brown. 2021. "Effects of an urban sanitation intervention on childhood enteric infection and diarrhea in Maputo, Mozambique: A controlled before-and-after trial." *eLife* 10:e62278. https://doi.org/10.7554/eLife.62278

Krause, Matthias, Enrique Cabrera Rochera, Francisco Cubillo, Carlos Díaz, and Jorge Ducci. 2018. "An International Standard for Assessing Water and Wastewater Services." AquaRating. Inter-American Development Bank (IDB).

Lambert, Donald. 2021. "Water Pressure: Eight Reforms to Boost Investment into Viet Nam's Water Sector | Asian Development Blog." *Asian Development Blog* (blog). July 5, 2021. https://blogs.adb.org/blog/water-pressure-eight-reforms-boost-investment-viet-nam-s-water-sector.

Manila Water. 2022. "2022 Integrated Report - Unwavering: 25 Years of Exceptional Service." https://www.manilawater.com/storage/files/1/investor/Results%20and%20Presentations/2022-Integrated-Report.pdf.

Mann, Robin S., Ahmed Abbas, Holger Kohl, Ronald Orth, and Mario Görmer. 2010. "Global Survey on Business Improvement and Benchmarking." Global Benchmarking Network. https://www.globalbenchmarking.org/wp-content/uploads/2018/07/2010_GBN_Survey_Business_Improvement_and_Benchmarking.pdf.

Maqbool, Nida, Muhammad Arslan Shahid, and Sher Jamal Khan. 2022. "Situational Assessment for Fecal Sludge Management in Major Cities of Pakistan." *Environmental Science and Pollution Research*, August. https://doi.org/10.1007/s11356-022-22331-2.

McDonald, David A. 2016. "The Weight of Water: Benchmarking for Public Water Services." *Environment and Planning A: Economy and Space* 48 (11): 2181–2200. https://doi.org/10.1177/0308518X16654913.

Mehta, M., D. Mehta, and A. Immanuel. 2011. "Benchmarking in Emerging Economies: The Performance Assessment System (PAS) Project in India." Center for Water and Sanitation, CEPT University. https://cwas.org.in/cwas-resources/benchmarking-in-emerging-economies-the-performance-assessment-system-pas-project-in-india.

Mehta, Meera, and Dinesh Mehta. 2014. "Water and Sanitation Service Levels in Cities of India (2011-12 and 2012-13)." PerformanceAssessmentSystem(PAS); CEPT University.

Minelli, Andrea. 2021. "Performance Benchmarking in the Water Utilities Sector," February 23, 2021. https://www.performancemagazine.org/benchmarking-water-utilities-sector/.

Ministry of Local Government, Rural Development and Co-operatives. 2017. "Institutional and Regulatory Framework for Fecal Sludge Management (IRF-FSM)." Bangladesh: Government of Bangladesh. https://psb.gov.bd/policies/fsmpe.pdf.

——— 2021. "Implementation of Institutional and Regulatory Framework for Fecal Sludge Management National Action Plan (NAP) 2021-2030." Bangladesh: Government of Bangladesh. https://www.psb.gov.bd/policies/iirffsme.pdf.

"Monitoring & Performance Reporting." n.d. National Water Supply & Sanitation Council. Accessed February 23, 2024. https://www.nwasco.org.zm/index.php/regulatory-tools/monitoring-performance-reporting.

Mugisha, Silver, Sanford V. Berg, and William T. Muhairwe. 2007. "Using Internal Incentive Contracts to Improve Water Utility Performance: The Case of Uganda's NWSC." SSRN Scholarly Paper. Rochester, NY. https://papers.ssrn.com/abstract=1088146.

Murungi, Caroline, and Maarten W. Blokland. 2016. "Benchmarking for the Provision of Water Supply and Sanitation Services to the Urban Poor: An Assessment Framework." *International Journal of Water* 10 (2–3): 155–74.

Muximpua, Odete, and Gaia Hatzfeldt. 2020. "One Single Amenity Can Make All the Difference to Gender Diversity in the Workplace." *World Bank Blogs* (blog). July 9, 2020. https://blogs.worldbank.org/water/one-single-amenity-can-make-all-difference-gender-diversity-workplace.

NWASCO. 2014. "Guidelines on Tariff Setting." https://www.nwasco.org.zm/files/Guidelines_on_Tariff_Setting_.pdf.

——— 2018. "Rural Water Supply and Sanitation - Framework for Provision and Regulation in Zambia." Lusaka, Zambia: National Water Supply and Sanitation Council of Zambia.

——— 2021. "Urban and Peri-Urban Water Supply and Sanitation." National Water Supply and Sanitation Council of Zambia. http://www.nwasco.org.zm/index.php/media-center/publications/urban-and-peri-urban-wss-sector-reports/send/12-urban-and-peri-urban-wss-sector-reports/81-urban-and-peri-urban-water-supply-and-sanitation-sector-report-2021.

——— 2022. "Water Supply and Sanitation Sector Report." https://www.nwasco.org.zm/index.php/media-center/publications/water-supply-and-sanitation-sector-reports?task=download.send&id=94&catid=12&m=0.

Nzitonda, Jacques. 2022. "Towards Improving Faecal Sludge Management in Kigali, Rwanda." IWA.

Ofwat. 2017. "Approach to Enforcement." Ofwat: The Water Service Regulation Authority.

——— 2022. "Water Company Performance Report 2021-22." Ofwat.

——— 2023. "Water Sector Overview." Ofwat. 2023. https://www.ofwat.gov.uk/regulated-companies/ofwat-industry-overview/.

OpenAI. 2021. "ChatGPT." 2021. https://chat.openai.com.

Peal, Andy, Barbara Evans, Isabel Blackett, Peter Hawkins, and Chris Heymans. 2015. "A Review of Fecal Sludge Management in 12 Cities." Final Report. Word Bank. https://www.susana.org/_resources/documents/default/3-2212-7-1435304068.pdf.

Pillay, Sudhir, and Jay Bhagwan. n.d. "Towards Establishment of the Brown Drop – A Regulatory Platform in Managing Excreta Flows in South Africa." Water Research Commission.

Practical Action. 2023. "Putting Pride into Poo." *Practical Action* (blog). 2023. https://practicalaction.org/our-work/projects/putting-pride-into-poo-faecal-sludge-management/.

PWWA. 2023. "Benchmarking." Pacific Water and Wastewater Association. 2023. https://www.pwwa.ws/benchmarking/.

Ricks, Jacob I., and Amy H. Liu. 2018. "Process-Tracing Research Designs: A Practical Guide." *PS: Political Science & Politics* 51 (4): 842–46. https://doi.org/10.1017/S1049096518000975.

Rivera Jr., Virgilio. 2013. "Manila Water Company, Inc: A Case Study on Private-Public Partnership." Manila Water Company, Inc.

Robbins, David, Linda Strande, and Julian Doczi. 2021. "Opportunities in Fecal Sludge Management for Cities in Developing Countries: Experiences from the Philippines." https://www.fsmtoolbox.com/assets/pdf/15.pdf.

RURA. 2021. "Rwanda Utilities Regulatory Authority_Annual Report_2020-2021." Annual Report.

"RUSH - Rural Urban Sanitation and Hygiene." n.d. Accessed February 23, 2024. https://rtmis.akvotest.org/.

Sanjeevi, V, and P Shahabudeen. 2015. "Development of Performance Indicators for Municipal Solid Waste Management (PIMS): A Review." *Waste Management & Research: The Journal for a Sustainable Circular Economy* 33 (12): 1052–65. https://doi.org/10.1177/0734242X15607428.

Sekayizzi, Andrew, Charles Odonga, and Johnson Amayo. 2005. "The National Water and Sewerage Corporation Experience in the Provision of Services to the Urban Poor," January. https://repository.lboro.ac.uk/articles/conference_contribution/The_National_Water_and_Sewerage_Corporation_experience_in_the_provision_of_services_to_the_urban_poor/9589556/1.

Senkwe, Barbara, and Charles B. Niwagaba. 2016. "Benchmarking Visit to eThekwini Municipality by Staff of Lusaka Water and Sewerage Company (LWSC), Zambia; Kampala Capital City Authority (KCCA) and the National Water and Sewerage Corporation (NWSC), Uganda." AfWA.

Setty, K, G O'Flaherty, J Enault, S Lapouge, JF Loret, and J Bartram. 2018. "Assessing Operational Performance Benefits of a Water Safety Plan Implemented in Southwestern France." *Perspectives in Public Health* 138 (5): 270–78. https://doi.org/10.1177/1757913918787846.

SFD Promotion Initiative. 2018. "SFD Manual (Volume 1 and 2, Version 2.0)." https://www.susana.org/_resources/documents/default/3-2357-7-1529046600.pdf#nameddest=Section_5.

Sharma, Anupam. 2006. "Urban Water Sector in South Asia: Benchmarking Performance." Field Note 37363. New Delhi, India: Water and Sanitation Program (WSP).

Sinha, Sunil. 2013. *Utility Benchmarking for Drinking Water: Synthesis Report*. IWA Publishing.

Sisson, Keith, James Arrowsmith, and Paul Marginson. 2003. "All Benchmarkers Now? Benchmarking and the 'Europeanisation' of Industrial Relations." *Industrial Relations Journal* 34 (1): 15–31. https://doi.org/10.1111/1468-2338.00256.

SNV Bangladesh. 2020. "IMIS: Improving Municipal-Led FSM Service Efficiency and Accountability." October 22, 2020. https://snv.org/update/imis-improving-municipal-led-fsm-service-efficiency-and-accountability.

——— 2021a. "Key Performance Indicators (KPIs) to Professionalise FSM Services in Bangladesh." REPORT (USHHD). Dhaka, Bangladesh.

——— 2021b. "Public Financial Management for FSM Services: A Gap Analysis and Some Recommendations." Urban. Dhaka.

Still, David, and Faye Balfour. 2006. "The Use of Key Performance Indicators in the Benchmarking of Rural Water Supply Schemes." Water Research Commission. https://www.wrc.org.za/wp-content/uploads/mdocs/TT255-06.pdf.

Strande, Linda, Barbara Evans, Marcos von Sperling, Jamie Bartram, Hidenori Harada, Anne Nakagiri, and Viet-Anh Nguyen. 2023. "Urban Sanitation: New Terminology for Globally Relevant Solutions?" *Environmental Science & Technology* 57 (42): 15771–79. https://doi.org/10.1021/acs.est.3c04431.

Strande, Linda, Mariska Ronteltap, and Damir Brdjanovic. 2014. *Faecal Sludge Management: Systems Approach for Implementation and Operation*. IWA Publishing. https://doi.org/10.2166/9781780404738.

Strande, Linda, Lars Schoebitz, Fabian Bischo, Daniel Ddiba, Francis Okello, Miriam Englund, Barbara J Ward, and Charles B Niwagaba. 2018. "Methods to Reliably Estimate Faecal Sludge Quantities and Qualities for the Design of Treatment Technologies and Management Solutions." *Journal of Environmental Management* 223 (June): 898–907. https://doi.org/10.1016/j.jenvman.2018.06.100.

Sustainable Sanitation Alliance. 2023. "SFD Data." 2023. https://sfd.susana.org/#.

The Aquaya Institute. 2021. "Data for Decision Making: Water and Sanitation in Low-Resource Settings (Annex I: Analysis of Results by Use Cases)." https://aquaya.org/wash-data-for-decision-making/.

——— 2022a. "Excreta Flow Diagram Dashboard." Tableau. https://public.tableau.com/app/profile/aquaya/viz/SHITFLOWDIAGRAMDASHBOARD/SHITFLOWDIAGRAMDASHBOARD.

——— 2022b. "Data for Decision-Making: Water and Sanitation in Low-Resource Settings."

Thiadens, Albert, and Ernest Betham. 2013. "Pacific Water and Wastewater Utilities Benchmarking Report." Pacific Water and Wastes Association; Asian Development Bank.

TNA. n.d. "Water Services Regulation Authority (OFWAT) Paper and Electronic Records 1989-2010." The National Archives.

Trimmer, John T., Hannah A. C. Lohman, Diana M. Byrne, Stephanie A. Houser, Fulgensio Jjuuko, David Katende, Noble Banadda, Assata Zerai, Daniel C. Miller, and Jeremy S. Guest. 2020. "Navigating Multidimensional Social–Ecological System Trade-Offs across Sanitation Alternatives in an Urban Informal Settlement." *Environmental Science & Technology* 54 (19): 12641–53. https://doi.org/10.1021/acs.est.0c03296.

Trimmer, John T., Daniel C. Miller, Diana M. Byrne, Hannah A. C. Lohman, Noble Banadda, Katherine Baylis, Sherri M. Cook, et al. 2020. "Re-Envisioning Sanitation As a Human-Derived Resource System." *Environmental Science & Technology* 54 (17): 10446–59. https://doi.org/10.1021/acs.est.0c03318.

Tsinda, Aime. 2020. "Guidelines for Faecal Sludge Management." Rwanda Utilities Regulatory Authority (RURA).

Ty, Patrick Lester. 2023. "Concessionaires Reporting and Benchmarking." Presented at the Watsan Regulatory Conference for South Asia, Kathmandu, Nepal, September 7.

United Nations Human Rights Council. 2017. "Report of the Special Rapporteur on the Human Rights to Safe Drinking Water and Sanitation." A/HRC/36/45. https://documents.un.org/doc/undoc/gen/g17/216/95/pdf/g1721695.pdf?token=Fm0TPcgAsmBI2dEOUv&fe=true.

United Nations. 2015. "Transforming Our World: The 2030 Agenda for Sustainable Development." https://sdgs.un.org/sites/default/files/publications/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf.

URBAN WASH. 2023. "Effective Enabling Environments for Inclusive Urban Water and Sanitation Provision." Washington, D.C.: USAID Urban Resilience by Building and Applying New Evidence in Water, Sanitation, and Hygiene (URBAN WASH). https://www.globalwaters.org/resources/assets/effective-enabling-environments-inclusive-urban-water-and-sanitation-provision.

United States Government. 2022. "U.S. Global Water Strategy 2022-2027."

USAID Learning Lab. 2022. "Theory of Change Workbook: A Step-by-Step Process for Developing or Strengthening Theories of Change." 2022. https://usaidlearninglab.org/resources/theory-change-workbook-step-step-process-developing-or-strengthening-theories-change.

Vavaliya, Jaladhi, and Dhruv Bhavsar. 2022. "Performance Assessment System for Evidence-Based Decision Making in Water and Sanitation Services: Implementation Experience and Lessons Learned." Stanford Woods Institute for the Environent.

Velidandla, Sasanka, Krishna C Rao, Reema Parekh, and Praveen Nagaraja. 2020. "Quality in Faecal Sludge Management: Benchmarks, Standards and Specifications." New Delhi: Water Sanitation and Hygiene Institute; NFSSM Alliance.

Wait, Suzanne, and Ellen Nolte. 2005. "Benchmarking Health Systems: Trends, Conceptual Issues and Future Perspectives." Edited by F Graham and Stewart Lawrence. *Benchmarking: An International Journal* 12 (5): 436–48. https://doi.org/10.1108/14635770510619366.

Waltz, Thomas J., Byron J. Powell, María E. Fernández, Brenton Abadie, and Laura J. Damschroder. 2019. "Choosing Implementation Strategies to Address Contextual Barriers: Diversity in Recommendations and Future Directions." *Implementation Science* 14 (1): 42. https://doi.org/10.1186/s13012-019-0892-4.

WASH-FIN. 2021. "A Financing Framework for Water Supply and Sanitation in the Philippines." Country Brief Series. USAID's Water, Sanitation and Hygiene Finance (WASH-FIN) Program. https://www.globalwaters.org/sites/default/files/wash-fin_program_country_brief_philippines_0.pdf.

WASREB. 2022. "A Performance Report of Kenya's Water Services Sector – 2020/21." Issue No 14. Impact. Nairobi, Kenya: WASREB. https://wasreb.go.ke/impact-report-issue-no-14/.

——— 2023. "Impact: A Performance Report of Kenya's Water Services Sector - 2021/22." Issue 15.

Weiner, Bryan J. 2009. "A Theory of Organizational Readiness for Change." *Implementation Science* 4 (1): 67. https://doi.org/10.1186/1748-5908-4-67.

WHO. 2021a. "Reflecting on TrackFin 2012–2020: Key Results, Lessons Learned and the Way Forward." World Health Organization. https://www.who.int/publications/i/item/9789240028432.

——— 2021b. "WHO Global Benchmarking Tool (GBT) for Evaluation of National Regulatory Systems of Medical Products, Revision VI." Geneva: World Health Organization. https://iris.who.int/bitstream/handle/10665/341243/9789240020245-eng.pdf?sequence=1.

WHO UNICEF Joint Monitoring Programme [JMP]. 2023. "WASH Data." 2023. https://washdata.org/data.

World Bank, dir. 2016. *Solar Pumping: A Cheaper and Cleaner Way to Access Groundwater*. https://www.youtube.com/watch?v=5yQOD1IBbjg.

——— 2019. "Women in Water Utilities: Breaking Barriers." Washington, D.C.: World Bank. https://www.susana.org/_resources/documents/default/3-3779-236-1582456002.pdf.

——— 2023. "Citywide Inclusive Sanitation (CWIS) Initiative." 2023. https://www.worldbank.org/en/topic/sanitation/brief/citywide-inclusive-sanitation.

WSP. 2018. "Benchmarking of Urban Water and Sanitation Facilities Gains Momentum in India." Water and Sanitation Program. 2018. https://www.wsp.org/featuresevents/features/benchmarking-urban-water-and-sanitation-facilities-gains-momentum-india.

WSUP. 2015. "Stand-Alone Unit or Mainstreamed Responsibility: How Can Water Utilities Serve Low-Income Customers?" Water & Sanitation for the Urban Poor. https://www.wsup.com/content/uploads/2017/08/DP005-Stand-alone-unit-or-mainstreamed-responsibility.pdf.

——— 2018. "Sanitation Set for Major Upgrade in Lusaka, Zambia."

——— 2019. "Systems Reboot: Sanitation Sector Change in Maputo and Lusaka." Water & Sanitation for the Urban Poor. https://www.wsup.com/content/uploads/2019/11/DP013-Systems-reboot-sanitation-sector-change-in-Maputo-and-Lusaka.pdf.

APPENDIX A: CASE STUDY SELECTION AND RESPONDENT AFFILIATIONS

List A1: Expanded interim list of case study locations identified using multiple data sources, prior to applying exclusion criteria

1. Bangladesh*
2. Benin^
3. Bolivia^
4. Burkina Faso^
5. Cambodia^
6. Cameroon^
7. Colombia^
8. Ethiopia^
9. Haiti**
10. India*
11. Indonesia**
12. Ghana^
13. Kenya**
14. Lesotho^
15. Madagascar^
16. Malaysia^
17. Mali^
18. Mozambique^
19. Nepal^
20. Niger^
21. Peru^
22. Philippines**
23. Rwanda#
24. Senegal^
25. Sierra Leone^
26. South Africa^
27. Tanzania#
28. Thailand^
29. Uganda*
30. Zambia*

*Short-listed from literature review and pursued for depth

**Added and pursued for breadth

#Dropped due to contact availability or lack of compelling evidence of benchmarking progress

Table A1: Variability in region and country income level for locations included in case study analysis

Country	Country Income	World Region
Uganda	Low	Africa
Zambia	Low	Africa
Bangladesh	Lower middle	Asia
India	Lower middle	Asia
Indonesia	Upper middle	Asia
Philippines	Lower middle	Asia
Kenya	Lower middle	Africa
Haiti	Lower middle	Latin America and Caribbean

Table A2: Key informant affiliations and actor types

Location	Affiliation	Sector	Role
Global	Fecal Sludge Management (FSM) Alliance	Network	Research and learning
	Sustainable Sanitation Alliance/German Development Agency (GIZ)	Network	Research and learning
	Asian Institute of Technology (AIT) Thailand, Global Water & Sanitation Center	Education	Implementation support
	Athena Infonomics	Private	Implementation support, research and learning
	Eawag	Education	Research and learning
Uganda	Ministry of Water & Environment	Public	Regulator
	National Water and Sewerage Corporation (NWSC)	Public	Service provider
	Kampala Capital City Authority (KCCA)	Public	Service provider
	Sheercare Services	Private	Service provider
	Makerere University	Education	Research and learning
Zambia	National Water Supply and Sanitation Council (NWASCO)	Public	Regulator
	Lusaka Water Supply and Sanitation Company (LWSC)	Public	Service provider
	Tiyende Sanicon Limited	Private	Service provider
	Water & Sanitation for the Urban Poor (WSUP)	Nongovernmental	Implementation support
Bangladesh	WaterAid	Nongovernmental	Implementation support
	SNV	Nongovernmental	Implementation support
India	CEPT University	Education	Implementation support, research and learning
	Indian Institute for Human Settlements	Education	Implementation support, research and learning
Philippines	Metropolitan Waterworks and Sewerage System Regulatory Office	Public	Regulator
Indonesia	Islamic Development Bank (isDB)	Donor	Finance
Kenya	Fresh Life	Nongovernmental	Service provider

Location	Affiliation	Sector	Role
Haiti	Sustainable Organic Integrated Livelihoods (SOIL)	Nongovernmental	Service provider

APPENDIX B: SEMI-STRUCTURED INTERVIEW GUIDE

Introduction: Thank you for agreeing to participate in this study. I am from The Aquaya Institute and am conducting this research on behalf of the Urban Resilience by Building and Applying New Evidence in Water, Sanitation, and Hygiene (URBAN WASH) consortium, which is led by Tetra Tech and sponsored by the United States Agency for International Development (USAID). The goal of this research is to understand best practices for tracking fecal sludge management (FSM) performance in different contexts.

We expect this interview to take up to 45 minutes and value your time and contributions! The results will be aggregated and not tied to individual contributors. You are welcome to skip any questions not applicable to your experience or add information that you feel is pertinent. Do you have any questions before we begin?

[If by phone:] Is it okay to record this call?

Demographics

1. What is your name, organization, and role?
2. How long have you served in this role?
    a. (If <1 year) Which other sanitation provision roles, if any, have you previously served in?

Referrals

1. If possible, please provide the name(s) and contact information of 2–3 colleagues who might be able to provide more information on FSM benchmarking practices.

FSM approaches

1. What portion of the population in the case study area uses on-site sanitation?
    a. Where is this information documented? If it is not publicly accessible, could you please share a copy?
    b. Which sub-populations or areas might be lacking access to sanitation services?
2. Are other FSM service providers active in the case study area?
    a. Which sub-populations do they serve?
3. How are the FSM services going in general? Are there any notable strengths or challenges? (Can use the table below to record notes.)

Services	General Approach	Strengths	Challenges
Coverage of full sanitation chain (i.e., containment, collection, transport, treatment, disposal, reuse)			
Governance			
Incentives (e.g., regulatory mandates)			
Management or implementation			
Technical aspects			
Financial health			
Social inclusion			

FSM benchmarking approaches (where benchmarking is in place)

We define "benchmarking" as assessing performance metrics using quantitative indicators that can be compared to targets, historical values, and other service providers. In addition, "process" benchmarking might involve improving practices through peer knowledge sharing and imitation (e.g., adopting specific protocols or guidance).

1. How would you define benchmarking?
2. Is FSM benchmarking taking place?
    a. If not, do you think introducing benchmarks would be helpful or not? In what ways and for whom? (can skip remaining questions in this section)
    b. How long has benchmarking been practiced?
    c. Who initiated the benchmarking efforts and why? For example:
        i. To promote and motivate competition among service providers
        ii. To identify strengths and weaknesses in service provider performance
        iii. To promote information sharing and improve transparency in reporting
        iv. To examine performance trends over time
        v. To provide information regarding to consumers
        vi. To improve data accuracy for global benchmarking efforts
    d. Which actors are responsible for collecting, validating, and submitting data on benchmarking indicators?
    e. To whom is benchmarking data submitted?
3. How is FSM benchmarking going in general? What are the strengths and challenges? Have any of these changed or been addressed over time?

	General Approach	Strengths	Challenges
Incentives (e.g., targets, funding)			
Data quality/verification			
Data reporting			
Data use			
Goal achievement			
Other			

4. Does FSM benchmarking cover all stages of the sanitation value chain equally well?
    a. Containment
    b. Collection/transport
    c. Treatment
    d. Disposal/reuse
5. Are key performance indicators (KPIs) used to support metric benchmarking?
    a. How many KPIs are monitored regularly?
        i. Where can we find a current listing?

b. How were the KPIs initially proposed and adopted?

c. Was/is there any resistance to adopting the KPIs?

d. Do any KPIs consider gender or social inclusion?

e. Do any KPIs address climate adaptation?

f. Which, if any, KPIs might be hard to measure or biased or otherwise create issues?

6. Are you aware of any other types of ongoing benchmarking (e.g., process comparison, third-party assessments)?

7. What are the future priorities, if any, for FSM benchmarking?

   a. How do the current KPIs reflect (or not reflect) the strategy to achieve these priorities?

   b. How do vulnerable groups (e.g., women and girls, people with disabilities, low-income area residents) fit into these priorities?

   c. Are any new KPIs proposed?

   d. Are any innovative approaches to benchmarking being piloted or anticipated?

8. What have been the positive or negative outcomes of benchmarking efforts to date?

   a. Are you aware of differences between your performance and that of your peers? If so, how?

   b. Are you aware of changes in your performance over time? If so, how?

   c. Would you say the objective(s) of the benchmarking program have been achieved? Why or why not?

   d. Have policy or management approaches been updated in response to benchmarking findings?

   e. Has benchmarking led to any other impacts?

APPENDIX C: CHRONOLOGY OF WATER AND WASTEWATER BENCHMARKING INITIATIVES

Table C1: Examples of domestic water and wastewater benchmarking initiatives, sorted in chronological order by start date

Country	Start Date	Benchmarking Approach (formal name, if applicable)	Coordinating or Lead Entity	Covered Services	Public Sharing of Results	Financial Incentives or Penalties
United Kingdom	1989	Regulatory monitoring	Water Services Regulation Authority (independent regulator)	Water and sewerage	Yes	Refund to customers for underperformance and increase payments for overperformance
Brazil	1996	Regulatory monitoring (National System for Water and Sanitation Data)	Ministry of Rural Development	Water and sewerage	Yes	Municipalities that fail to provide the information cannot obtain federal funds
Canada	1997	Utility-led monitoring (Canadian Infrastructure Benchmarking Initiative)	AECOM (private consulting firm)	Water	No, considered "commercial confidential"	None
Netherlands	1997–2011; 2012–present	Utility-led monitoring; regulatory monitoring	Vewin (Association of Dutch Water Companies)	Water and sewerage	Yes	None
Zambia	2000	Regulatory monitoring	National Water Supply and Sanitation Council (NWASCO; independent regulator)	Water and on-site sanitation	Yes	Penalties can include enforcement notices, fines, and in worst-case scenarios, withdrawal of their license
India	2003	Regulatory monitoring (Service-Level Benchmarks)	Ministry of Urban Development and Urban Local Bodies (municipal governments)	Water and sewerage	Yes	States that take part in benchmarking are eligible for grants from the Central Finance Commission
Indonesia	2003	Utility-led monitoring	Indonesian Association of Water Supply Companies	Water	Encourage sharing at the utility level	Provides subsidies to service providers involved in benchmarking
Australia	2004	Regulatory monitoring	National Water Commission (central	Water	Yes	None

Country	Start Date	Benchmarking Approach (formal name, if applicable)	Coordinating or Lead Entity	Covered Services	Public Sharing of Results	Financial Incentives or Penalties
		(National Water Initiative)	government agency)			
Bangladesh	2005	Regulatory monitoring	Ministry of Local Government, Rural Development, and Cooperatives	Water	Yes	None
Pakistan	2005	Regulatory monitoring	National government	Water and sewerage	Yes	None
South Africa	2008	Utility-led monitoring (Municipal Benchmarking Initiative)	South Africa Local Government Association	Water and on-site sanitation	By category but not by provider	None, voluntary initiative
Portugal	2009	Regulatory monitoring	Water and Waste Services Regulation Authority (independent regulator)	Water and sewerage	Yes	Awards and sanctions
Zimbabwe	2012	Regulatory monitoring	Local authorities and the Peer Review Coordinating Committee	Water and sewerage	Yes	Benchmarking data informs public-sector investment programs
Tanzania	2014	Regulatory monitoring	Energy and Water Utilities Regulatory Authority (EWURA) (independent regulator)	Water and sewerage	Yes	Best-performing utilities receive financial rewards
Kenya	2015	Regulatory monitoring	Water Services Regulatory Board (WASREB; independent regulator)	Water and sewerage	Yes	Incentives for good performance. Sanction for poor performance can include withdrawal of license

Table C2: Examples of water benchmarking activities embedded in performance-based contracts in the context of public-private partnerships, sorted in chronological order by start date

Location	Start Date	Service Provider(s) being Benchmarked	Oversight
Burkina Faso	1993	Various (e.g., 2001–2008 contract was awarded to the consortium of *Compagnie Générale des Eaux and Mazard et Guérard* [now Veolia, France])	National Office of Water and Sanitation (ONEA)
Senegal	1996	1996–2019 : *Sénégalèse Des Eaux* ; 2020–present : Suez	National Water Company of Senegal (SONES)
Columbia	1996	Combination of large international operators and small and medium national groups	Superintendency of Residential Public Services (SSP)
Manila, Philippines	1997	Manila Water	Municipal government
Jakarta, Indonesia	1997	Permuda Air Minum Jaya	Municipal government
Uganda	1998	National Water and Sewerage Corporation (NWSC) branches	NWSC headquarters
Morocco (Casablanca, Rabat, Tangiers, and Tetouan)	1999 (Casablanca)	Various examples include Lydec (Casablanca), Amendis (Tangier and Tetouan), and Redal (Rabat)	National Office of Potable Water (ONEP)
Niger	2001	Various (e.g., 56 private companies operating in rural areas)	Niger Water Public Asset Holding Company (SPEN) and Niger Water Operating Company (SEEN)
Malaysia (Sabah)	2003	Halcrow Water Services in partnership with a Malaysian company, Salcon Engineering	Sabah Water Board

APPENDIX D: KEY PERFORMANCE INDICATORS

Table D1: Examples of operational, financial, and organizational KPIs used in the water and wastewater sectors

Type	KPI	Definition	Limitations
Operational	Service coverage	Percentage of the population in service area connected to the water/wastewater network	Not indicative of service coverage in low-income areas located outside the official service area
Operational	Metering ratio	Percentage of (legal) water connections with an active meter	Does not encourage efforts to identify illegal connections
Operational	Reliability/continuity of water service	Hours of availability per day	Does not indicate time of day (and therefore ease of access)
Operational	Non-revenue water	Volume of produced water that does not result in revenue due to physical losses (e.g., leaks) or commercial losses (e.g., billing inaccuracies, customers defaulting on payment or exempt from paying), expressed as a percentage or volume supplied per connection per day	Does not always capture diverse causes of high non-revenue water; lack of consensus on the best definition (percentage or volume per connection per day), making comparisons difficult (AWWA 2019).
Operational	Customer complaints	Number of complaints per connection (percentage), response time to complaints, or percentage of complaints adequately resolved	Not commonly tracked; improved monitoring may initially appear to increase rates
Operational	Energy use	Kilowatts of electricity per cubic meter of produced water or sewerage	—
Operational	Sewer system blockages	Blockages per kilometer of network per year	Not commonly tracked; improved monitoring may initially appear to increase rates
Operational	Level of sewage treatment	Levels of primary and secondary treatment	Does not capture later treatment steps that address microbiological contamination
Operational	Pipe breaks	Pipe breaks per kilometer of network per year	Not commonly tracked; improved monitoring may initially appear to increase rates
Operational	Water mains rehabilitation	Percentage of water mains rehabilitated per year	—
Operational	Training	Number of training days per staff per year	—
Operational	Maintenance	Existence of a maintenance plan	—
Operational	Microbiology compliance	Percentage of water samples compliant with microbiological water quality standards	Different service providers may use different denominators (e.g., required number of samples versus total number of collected

Type	KPI	Definition	Limitations
			samples), complicating comparisons
Operational	Chlorine residual compliance	Number of chlorine residual tests conducted compared to the number of tests required by applicable standards (percentage)	Could use total chlorine or free chlorine for water distribution system
Operational	Service coverage in low-income areas	Percentage of population in low-income areas served by the water or wastewater network	Difficult to quantify due to fast-evolving low-income area locations and population sizes; may obscure that service modalities are different in low-income areas compared to other parts of the service area (e.g., delegation of services to informal providers or a community-based organization to minimize financial risks)
Financial	Operating cost coverage ratio (or operating cost recovery)	Ratio of revenue to operation and maintenance costs	Does not capture true economic costs since it excludes capital improvements, which no KPI explicitly captures (although increases in service coverage and continuity usually result from capital improvements)
Financial	Bill collection efficiency	Percentage of bills that are paid or collected	Does not distinguish among causes of low efficiency (e.g., household ability to pay, utility collection practices)
Financial	Affordability	Tariff per capita per year as a percentage of the country's annual gross national income per capita	Using national income as the denominator overlooks regional disparities
Financial	Debt	Amount of debt payments per year or debt ratio (percentage of expenses for debt servicing)	—
Organizational	Staff productivity	Number of staff per 1,000 connections	Difficult to measure when services are outsourced to a contractor (e.g., number of contract staff and their hours)
Organizational	Staff motivation	Presence of a reward or recognition program for staff	—
Organizational	Female representation	Percentage of female staff, managers, or engineers	Does not address gendered wage gaps
Organizational	Female compensation	Women's average salary	Unless disaggregated by job type, does not allow distinguishing gendered wage gaps from low female representation in senior positions

Table D2: Examples of KPIs used for fecal sludge management (FSM), considering the literature review, case studies, and an elicitation exercise among water, sanitation, and hygiene (WASH) professionals (*prioritized indicators reflected in Table 6 appear in shaded rows)

Categories	KPI	Definition or Measurement Guidance	Source	Pros	Cons	Priority*
Access	Percentage of population with access to safely managed sanitation (household, health facilities, schools)	Summary of population coverage, meeting multiple criteria across the sanitation service chain (use of an improved sanitation facility that is not shared with other households and where excreta is safely disposed in situ or excreta is removed and treated offsite)	Joint Monitoring Programme (JMP); Citywide Inclusive Sanitation (CWIS); African Sanitation Policy Guidelines (AMCOW 2021); Eastern and Southern Africa Water and Sanitation Regulators Association (ESAWAS) 2022a; Tsinda 2020 (Rwanda); Key informants (Kenya)	Conforms with global monitoring priorities	May be difficult to measure all components accurately	High
Access, Equity	Ratio of low-income population with access to safely managed sanitation	Percentage of safely managed sanitation in low-income communities divided by percentage citywide	CWIS	Composite indicator; pro-poor	Income data, inventories, or maps of low-income areas likely to change frequently, some neighborhoods may not be homogenous	Medium
Access, Equity	Access to on-site sanitation services among vulnerable groups (e.g., transient populations)	Disaggregate, where relevant, by income, sex, age, race, ethnicity, migratory status, disability, and geographic location, in accordance with the fundamental principles of official statistics (AMCOW 2021, 119). United Nations (UN) Special Rapporteur on Water and Sanitation 2017 adds that equality is an integral part of the definition and disaggregation should include socioeconomic characteristics. Options (Box 24) include poorest versus richest wealth quintile, rural versus urban, informal settlements versus formal, disadvantaged groups versus general population	African Sanitation Policy Guidelines; UN Special Rapporteur on Water and Sanitation 2017	Focused on equitable access, as opposed to majority access	Unclear how to best identify and measure for all underserved populations, requires local adaptation	Medium
Containment, Collection	Percentage of emptiable facilities	Percentage of emptiable latrines, out of total latrines; Could measure via manual desludging prevalence or facility	EWURA 2020; Similar to Tsinda 2020 and ESAWAS 2022a	Assesses feasibility of offering on-site sanitation services	May be hard to collect data if not on-site, challenging to	Medium

Categories	KPI	Definition or Measurement Guidance	Source	Pros	Cons	Priority*
		type (e.g., pit latrine, septic tank, dry toilet, urine-diverting, pour-flush, Ecosan)			categorize consistently as new toilet styles become available	
Containment, Collection	Population using emptiable facilities	Percentage of emptiable latrines, out of total latrines, multiplied by average household size	ESAWAS 2022a; EWURA 2020	Assesses customer base	Have to estimate number of people using latrines	Low
Containment, Equity	Percentage of (wheelchair) accessible facilities	Percentage of the facilities that meet accessibility codes	Bangladesh "FS & SWM Dashboard," n.d.	Values leaving no one behind; disability can arise at any time	Accessibility standards may not exist in all countries	Medium
Containment, Equity	Access to safe toilets among low-income households	Percentage of low-income community population with access to safe individual toilets divided by percentage of total population with access to safe individual toilets	CWIS	Composite indicator; pro-poor	Hard to measure, unclear how they define "safe"	Low
Containment, Finance	Toilet sales cost or revenue	Average sales costs or revenue per toilet per month	Key informants (Haiti)	Potential income stream to supplement emptying fees, if offered by the service provider; potentially indicative of an increasing customer base	May be heavily subsidized, resulting in a net gain or loss; no guarantee that toilets will be used as planned	Low
Containment	Coverage of on-site sanitation facilities	Percentage of the population using pit latrines and septic tanks (versus shared public toilets, no toilets, sewered toilets) out of total population	Jayathilake et al. 2019; Performance Assessment System (PAS) in India (CEPT University 2015); shit-flow diagrams	Potentially compatible with JMP, could use a digital survey to measure, inventory forms a basis for other indicator assessments	Obtaining representative, up-to-date household-level data is resource-intensive, may come from infrequent one-time surveys	High
Containment	Percentage of households with access to improved toilets	E.g., pour-flush toilets with solid walls and doors, as opposed to pit latrines without a slab or platform, hanging latrines, or bucket latrines	JMP ladder; Bangladesh "FS & SWM Dashboard," n.d.; Tsinda 2020	Associated with better quality of life (Knee et al., 2021.); may capture consumer satisfaction	May distract from basic access and water conservation goals; quality of life standards may continue to shift	Medium
Containment	Percentage of on-site sanitation facilities conforming to standards	Percentage of households connected to septic tank as per design standards; alternatively, percentage of the facilities that meet construction standards	CEPT University 2015; Velidandla et al. 2020	Could bolster regulatory enforcement	Homeowners may be hesitant to report; standards for on-site sanitation may be under development	Medium

Categories	KPI	Definition or Measurement Guidance	Source	Pros	Cons	Priority*
	(construction codes)					
Containment	Percentage of public institutions (e.g., schools, hospitals) with on-site sanitation access, emptiable facilities, conforming to standards, or with improved toilets	Percentage of public spaces that have adequate sanitation facilities and where sludge is safely transported or disposed in situ	CWIS	Public toilets serve a greater segment of the population (sanitation justice); indicator tailoring might depend on local goals	May distract from household access	Medium
Containment	Number of toilet installations	Number of emptiable toilets installed at households per year	Key informants (Haiti)	Multiple case studies had unmet toilet installation needs to bridge the gap toward providing sustainable FSM services	The number alone does not reflect the remaining unmet need	Low
Containment	Willingness to pay for toilet improvement	(Variable)	Bangladesh "FS & SWM Dashboard," n.d.	Indirect association with financial viability of FSM	The number alone does not reflect the remaining unmet need	Low
Containment	Percentage of solid waste contaminating fecal sludge	Want to reduce and eliminate solid waste disposal in pit latrines because it is expensive to handle and reduces treatment options	Kampala Capital City Authority (KCCA) Strategy	Easy to measure (visual); relates to treatment cost	May be a location-specific issue	Low
Containment	Percentage of population practicing open defecation	Disposal of human feces in fields, forests, bushes, open bodies of water, beaches, and other open spaces or with solid waste	JMP ladder; Tsinda 2020	Still relevant, even in communities declared open defecation free	May reflect older global goals at the bottom of the service ladder; changes between night and day (reported in Uganda)	Low
Collection, Equity	Service area coverage (penetration ratio)	Percent of population or spatial area that has access to desludging services (In Bangladesh, the total number of containment facilities reached in last 5 years, divided by the total number of containments that can be accessed by a formal emptying service; in Rwanda, the number of districts with sanitation centers in place and operational)	Velidandla et al. 2020; SNV Bangladesh 2021a; Tsinda 2020	Calculates the reach of FSM services against the potentially reachable, which could help to identify the conversion rate from manual (or non-emptied) to mechanical emptying services	May be difficult to accurately estimate the "reachable" population; requires accurate maps	Medium

Categories	KPI	Definition or Measurement Guidance	Source	Pros	Cons	Priority*
Collection, Equity	Fecal sludge collection ratio in low-income areas	Volume of collected fecal sludge as a percentage of total estimated volume generated in the low-income service area	Staff exercise	Targets urban poor	Maps of low-income areas likely to change frequently; some neighborhoods may not be homogenous	Low
Collection, Equity	Inclusive service coverage in low-income areas	Percent of population or spatial area within low-income areas that has access to services (in Bangladesh, percentage of delivered services in low-income areas, out of total delivered services)	SNV Bangladesh 2021a	Targets urban poor, potential to approximate given the type of service providers operating (e.g., vacuum trucks vs. manual emptiers)	Maps of low-income areas likely to change frequently, some neighborhoods may not be homogenous	Low
Collection, Finance	Cost of emptying to service provider	Average annual direct and indirect emptying costs per toilet per month	Key informants (Haiti)	Relevant to operational cost recovery and efficiency; could inform service zones	May be subject to other influences (e.g., location of disposal points, number of trucks in fleet, fuel costs)	Low
Collection, Management	Customer satisfaction with emptying service	75% or higher rating on single-question survey (includes pricing satisfaction)	Velidandla et al. 2020; SNV Bangladesh 2021a	Measure of value to consumers; may have direct impacts on willingness to pay	Requires representative survey methods or baseline data to accurately assess changes	Medium
Collection, Management	Customer complaints	Number of customer complaints per month or per year (i.e., New International Benchmarking Network for Water and Sanitation Utilities [IBNET] tracks the number of complaints resolved; ESAWAS tracks billing complaints only)	New IBNET; ESAWAS 2022a; key informants (Philippines)	May be a valuable source of insight for regulators managing service providers; may already be tracked	Negative framing relative to customer satisfaction; hard to determine if data is representative if collected passively	Medium
Collection, Management	Response time to complaints	Percentage of complaints resolved within 48 hours or 5 days (for ESAWAS billing complaints only)	ESAWAS 2022a; Velidandla et al. 2020; Tsinda 2020; Ty 2023	Related to professional service perceptions and customer demand, could motivate a culture of performance improvement	Does not capture satisfaction with complaint resolution	Medium
Collection, Management	Ratio of emptiers to facilities	Number of registered emptiers per 1,000 tanks or pits	EWURA 2020	Addresses staffing sufficiency	May be restricted by contract agreements over multi-year periods	Low
Collection, Management	Customer retention rate	Average percentage of customers using the service year over year	Key informants (Haiti)	May provide early warning of service	Competition among service areas may be limited by	Low

Categories	KPI	Definition or Measurement Guidance	Source	Pros	Cons	Priority*
				issues and track business growth potential	contractual agreements in different contexts	
Collection, Management	Collection efficiency	Percentage of sludge collected out of the total expected to be collected (may be the same as the percentage of facilities emptied in practice)	CEPT University 2015; Jayathilake et al. 2019; EWURA 2020	Measures collection capacity and operating efficiency	Expected amounts would have to be estimated	Low
Collection, Management	Emptying technology	Number and types of mechanical and semi-mechanical technologies	Tsinda 2020	Related to emptying capacity	May change infrequently; may depend on age of fleet	Low
Collection, Treatment	Number of beneficiaries	Number of people benefitting from sludge collection and treatment services	Key informants (Zambia)	May resonate with donors and attract finance	Requires estimation; would be approximate at best	Low
Collection	Total sludge collection volume	Volume of fecal sludge collected per day	KCCA Strategy; Tsinda 2020; Key informants (Zambia)	Easy to measure, may better reflect population size served than number of facilities emptied	Hard to compare across locations with different population densities and mixes of sanitation services	High
Collection	Fecal sludge collection ratio	Volume of collected fecal sludge as a percentage of total estimated volume generated in the service area (in Bangladesh, the total amount of fecal sludge emptied from containments divided by the total estimated sludge to be emptied in accessible areas [calculated using accumulation rates])	ESAWAS 2022a; KCCA Strategy; SNV Bangladesh 2021a; Tsinda 2020	In theory, captures needs over the full service area	Denominator needs to be estimated and may be hard to validate; mix of septic tanks and unlined pit latrines, as well as variations in usage, make accurate volume tracking challenging (Gudda et al. 2019); requires desludging operators to submit data	Medium
Collection	Response time to requested desludging	Percentage of requests serviced within 48 hours (India) or average time between application and delivered service (Bangladesh)	Velidandla et al. 2020; SNV Bangladesh 2021a	May be easy to measure by comparing existing records	May need centralized call centers or websites (ticket system) to facilitate tracking	Medium
Collection	Emptying rate	Percentage of on-site sanitation facilities (pits or tanks) emptied per year out of the total number of on-site sanitation facilities (can be broken down by method)	ESAWAS 2022a; key informants (Zambia); Tsinda 2020; CEPT University 2015; EWURA 2020	Could help determine whether desludging needs to be scheduled comprehensively	Need on-site facility records; may have a mix of pit/tank designs and sizes with different emptying frequency recommendations	Medium
Collection	Number of facilities desludged	Total number of facilities (pits or tanks) desludged	Key informants (Philippines)	Easy to measure; related to frequency of emptying, which may be a carbon	Hard to compare across locations with different population densities and mixes of sanitation services	Medium

Categories	KPI	Definition or Measurement Guidance	Source	Pros	Cons	Priority*
				emission reduction goal		
Collection	Manual desludging prevalence	Percentage of fecal sludge removed manually as opposed to using a vacuum truck (conversely, percentage of desludging services completed mechanically or semi-mechanically); similar to the percentage of emptiable facilities	Bangladesh "FS & SWM Dashboard" n.d.; Tsinda 2020; EWURA 2020	May already be tracked by emptiers	Requires desludging operators to submit data, which may be illegal in some places, leading to under-reporting	Medium
Collection	Response efficiency	Percentage of services delivered, out of total applications (requests for service)	SNV Bangladesh 2021a	Clarifies the number of customers the service provider is potentially losing due to inattention, lack of capacity, inadequate access, etc.	Only relevant in locations where demand exceeds supply of services	Low
Collection, Social	Worker safety (during desludging)	Prevalence of injuries or deaths due to avoidable causes	Velidandla et al. 2020	Considers social impacts on workers	Under-reporting; requires monitoring of multiple work locations; may be chronic health effects that are hard to track; requires health or workers' compensation records	Low
Collection	Number of sludge collection operators	Number of emptiers active (e.g., at city level)	Staff exercise	May be indicative of consumer choice and market competition	Likely controlled by regulations (may not show changes over time)	Low
Collection	Number of vacuum-emptying machines	Ratio of septage-sucking machines for cesspit emptying per 1,000 septic tanks	ESAWAS 2022a	Indicates equipment capacity to support collection	Values may be estimated if accurate inventories are not available	Low
Transport, Environment	Frequency of safe transport of fecal sludge	Spillage and/or illegal dumping events per year or per 10,000 trips	Velidandla et al. 2020	Related to public health, source water protection	May be underreported; may require automated surveillance of truck weights	Medium
Transport, Environment	Number of dumping points on the sewer network	Aiming to increase from zero to ten	KCCA Strategy	Easy to measure; Related to transport time, safe disposal, and environmental outcomes	May be location-specific	Low
Transport	Rate of safe transport to designated points	Volume delivered to designated points, divided by volume collected	ESAWAS 2022a; Bangladesh "FS & SWM Dashboard" n.d.;	Related to public health and source water protection	May replicate volume collected; illegal dumping may not be reported; does not	High

Categories	KPI	Definition or Measurement Guidance	Source	Pros	Cons	Priority*
			Velidandla et al. 2020; SNV Bangladesh 2021a; Tsinda 2020; EWURA 2020		directly account for downstream disposal impacts	
Transport	Total volume delivered to designated points	Usually the same as volume collected	Key informants (Zambia); Tsinda 2020	May already be tracked at dumping sites (needed to calculate tipping fee)	May replicate volume collected	Medium
Transport	Desludging vehicle maintenance	Percentage of desludging vehicles that comply with maintenance standards; could measure via inspection and maintenance service regularity (e.g., every 6 months)	CWIS	Loosely related to regular inspections and maintenance, which might be more relevant	Vehicle conditions likely to vary among contexts; would be biased for an older fleet	Low
Treatment, Environment	Treatment ratio	Volume of properly treated sludge (both liquids and solids) as a percentage of the volume collected	ESAWAS 2022a; EWURA 2020	Indicative of progress along sanitation service chain and environmental and public health protection	Proper treatment likely to differ depending on regulatory context	High
Treatment, Equity	Worker safety (during treatment)	Prevalence of injuries or deaths due to avoidable causes (recommend splitting into measurable parts)	Velidandla et al. 2020	Considers social impacts on workers	Requires health or workers' compensation records; worker safety may be a crosscutting issue and deserve broader attention depending on the local context	Low
Treatment, Finance	Gate fees collected	Amount paid per week/month/year at treatment facility for fecal sludge disposal	Key informants (Zambia)	Data already available; proxy for treatment utilization and volume	May not vary substantially over time; other factors such as contracts may affect fee collection	Low
Treatment, Finance	Cost of sludge treatment	Average direct and indirect cost of treatment per toilet per month	Key informants (Haiti)	Indirectly related to treatment capacity, demand, and engineering options	May be difficult to separate costs for mixed sewerage and sludge treatment plants	Low
Treatment, Management	Treatment facility utilization	Ratio of volume treated to treatment capacity (in Bangladesh, annual fecal sludge volume treated divided by design capacity of fecal sludge treatment plant)	SNV Bangladesh 2021a	Related to long-term infrastructure, demand creation and service expansion planning	May depend on the types of waste (e.g., watery versus solids) and mixing processes in trucks or at treatment plants	Medium

Categories	KPI	Definition or Measurement Guidance	Source	Pros	Cons	Priority*
Treatment, Management	Sludge treatment process certification	Presence of certification mechanism for which treated fecal sludge has to qualify	CWIS	Aids international standardization	May carry an extra cost (e.g., to purchase International Organization for Standardization [ISO] standard and pay consultants)	Medium
Treatment, Management	Average audit score of treatment facilities	Categorical (e.g., high, medium, low) or numeric scale score of treatment plant's compliance with operational capacity expectations	Staff exercise	Summarizes achievement of multiple goals (e.g., treatment quality, environmental protection, worker safety)	More intensive data collection	Low
Treatment	Quality of treated fecal sludge	Percentage of tests that meet standards for biosolids, effluent, emissions, and proper waste disposal	ESAWAS 2022a; CEPT University 2015; Jayathilake et al. 2019; Velidandla et al. 2020; SNV Bangladesh 2021a; Tsinda 2020; EWURA 2020	Could reinforce reselling of sludge products, environmental and public health protection	Requires local field and/or laboratory testing capacity and standardized methods (external accredited labs preferred)	Medium
Treatment	Total volume of sludge treated	Volume of sludge processed by treatment facilities	Tsinda 2020	Easy to track volumes from tipping fee records	May depend on the types of waste (e.g., watery versus solids) and mixing processes in trucks or at treatment plant	Medium
Treatment	Treatment capacity	Volume of sludge that can be treated at city level (engineering capacity) or treatment capacity as a percentage of volume received	ESAWAS 2022a; CEPT University 2015; Jayathilake et al. 2019; KCCA Strategy; EWURA 2020	Helps to explain potential gaps in treatment rates	Unlikely to change over time	Medium
Treatment	Quality of water effluent	Percentage of compliant water quality parameters (e.g., biochemical oxygen demand [BOD], nitrate, phosphate, total suspended solids [TSS], total coliform, temperature)	SNV Bangladesh 2021a; Tsinda 2020; EWURA 2020	Ensures compliance with environmental regulations	Requires field and/or laboratory testing capabilities (external accredited labs preferred); may overlap with regulator responsibilities; may be more heavily influenced by sewerage, depending on facility type	Low
Treatment	Time spent per truck (in addition to active decanting) at treatment facility	Waiting time (hours/minutes) at treatment facility before and after active decanting	Velidandla et al. 2020	May relate to work efficiency, worker satisfaction, profit potential	May have to record or automate recording of check-in/check-out times to assess accurately	Low

Categories	KPI	Definition or Measurement Guidance	Source	Pros	Cons	Priority*
Disposal, Environment	Water contamination compliance	Fecal coliform level measured in receiving waters near treatment plant outfall	Bangladesh "FS & SWM Dashboard" n.d.	Encourages environmental stewardship and protection of wildlife, tourism	May be affected by mixing with sewerage; Priority may depend on usage of water body (e.g., fishing, recreation, drinking water)	Low
Disposal, Reuse	Percentage of fecal sludge safely managed	Shit-flow diagram: Combination of service delivery outcomes (e.g., sludge contained on-site or transported and treated), which results in hazard (excreta in the environment) and population exposure likely to result in a low public health risk	African Sanitation Policy Guidelines (2021); SFD Promotion Initiative 2018; Key informants	Targets population safety as well as environmental protection; when used within an SFD, the highly visual format can be useful as an advocacy and communication tool	Data intensive and takes a long time to update (cannot be completed frequently)	Medium
Disposal	Use of soak pits	Percentage of septic tanks connected to a soak pit for effluent disposal	ESAWAS 2022a	Allows an option for decentralized on-site disposal	Soak pits may have environmental and health drawbacks if not working properly	Low
Reuse, Environment	Fecal sludge reuse ratio	Volume of reused sludge (solids) as a percentage of the total volume treated	CEPT University 2015; Jayathilake et al. 2019; Velidandla et al. 2020; CWIS; EWURA 2020	Related to circular economy; neglected indicator category to date; data may already exist	Smaller activity scale relative to disposed waste; many countries lack laboratory testing capacity or standards to confirm fecal sludge safety for reuse; may not yet be adequately supported by policy and research sectors	Medium
Reuse, Environment	Wastewater reuse ratio	Volume of treated wastewater safely reused or discharged from fecal sludge treatment plant out of total treated effluent	ESAWAS 2022a; Velidandla et al. 2020; Tsinda 2020	May be easy to measure flow from point source(s)	May need costly additional infrastructure to facilitate non-potable water recycling; does not consider potential impacts to downstream sources from discharge to environment	Low
Reuse, Finance	Sales of sludge reuse products	Revenue collected from sale of reuse products or quantity and types of recovered products	Key informants (Haiti); Tsinda 2020	Data likely already exists, potentially related to carbon emissions (e.g., for briquettes)	Some volumes might be produced but not sold depending on the local market; may not be adequately supported by policy and research sectors	Medium

Categories	KPI	Definition or Measurement Guidance	Source	Pros	Cons	Priority*
Social, Collection, Treatment	Worker safety	Whether procedures and/or materials are in place and monitored for desludging and treatment plant operators and whether the government funds regular health checks	CWIS	Broad catch-all indicator for worker safety	May be disincentives for workers and employers to report safety violations in the absence of confidentiality protocols, external requirements, or audits; assumes adequate budgets	Low
Social, Equity	Gender pay gap in on-site sanitation workforce	Difference between median salary of male and female employees performing the same role	CWIS	Reflects incentive for women to participate	May be difficult to standardize across multiple participating organizations; may be biased due to manual labor demands of some positions	Low
Social, Equity	Percentage of female collection and treatment operators	Number of female workers divided by total number of workers	Staff exercise	Easy to measure	May reflect broader cultural influences outside of the FSM sector	Low
Social, Management	Female participation in decision-making workforce	Percentage of females in decision-making positions divided by total number of decision-making positions; CWIS suggests decision-making bodies are specifically government institutions; New IBNET tracks the overall percentage of female employees	CWIS/New IBNET	Easy to measure representation in decision-making positions	May not reflect influence on decision-making in practice	Medium
Social, Management	Sanitation worker equity (formalization, legal recourse, right to unionize, social security and health insurance)	Presence/absence of each indicator; IBNET/CWIS tracks whether labor legislation created for sanitation workers	CWIS/IBNET	Easy to monitor, not likely to change every year	Spectrum of what is expected may change over time relative to all locations; may get outdated easily; requires legal consultation	Low
Social	Worker safety procedures followed	Percentage of time adherence to guidelines and procedures observed over total observation time (in Bangladesh, percentage of services delivered following occupational health and safety guidelines, out of total amount of services delivered); IBNET/CWIS: Number and percentage of sanitation workers following standard operating procedures/health and safety standards in place and	IBNET; CWIS; SNV Bangladesh 2021a	More accurate measure of safety in practice than existence of procedures and materials	Requires third-party observation; difficult to observe consistently; may be disincentives to reporting	Medium

Categories	KPI	Definition or Measurement Guidance	Source	Pros	Cons	Priority*
		monitored and enforced and government-funded regular health checkups				
Social	Worker safety materials (personal protective equipment) available	Percentage of time adherence to equipment availability observed over total observation time	Staff exercise	Simple to observe	Requires third-party observation; may be affected by supply chains; represents intent rather than actual usage	Medium
Equity, Finance	Affordability (average household expenditures for on-site sanitation services)	Percentage of household expenditures spent on desludging costs, disaggregated across wealth quintiles	African Sanitation Policy Guidelines (Annex 1.2); Velidandla et al. 2020	Could aid tariff adjustment	Needs to be disaggregated to be useful; might lag behind present conditions in unstable economies; requires household income data; methods to estimate wealth vary	Medium
Equity	Public participation	Number of opportunities per year for public comment	African Sanitation Policy Guidelines (Annex 1.2)	Addresses gap noted in the case study exercise	Participation may be easy to measure, but meaningful or influential participation may be harder	Medium
Equity	Fecal-oral disease incidence	Rate of new disease cases over time	CWIS	Proxy for health and well-being, related to socioeconomic outcomes and health equity; requires good health records	Health outcomes are hard to measure in isolation of other influences; may discourage progress if they cannot be detected	Low
Finance, Collection	Average tariff or pit-emptying fee	Average emptying fee charged to household customers	CWIS/IBNET	Easy to track	Not as relevant as cost-recovery to financial health, nor as relevant as affordability to consumer well-being	Low
Finance, Collection	Customer conversion	Acceptance rate of offers for emptying	Key informants (Philippines)	Helps to inform sanitation marketing	Relevance depends on local market conditions	Low
Finance, Equity	Subsidy ratio allocated for on-site sanitation	Subsidy amount paid for non-sewered sanitation divided by the amount for sewered sanitation	CWIS	Easy to calculate, likely to be important during transition periods as service providers expand offerings	Interpretation may be context-specific as service arrangements are likely to vary; requires sewerage budget data	Medium
Finance, Management	Percentage of service providers	Percentage of contracts with payments allotted for specific performance	Staff exercise	Incentivizes performance-based funding	Performance-based funding does not work in all conditions	Low

Categories	KPI	Definition or Measurement Guidance	Source	Pros	Cons	Priority*
	with performance-based contracts	conditions out of total number of contracts				
Finance, Management	Strategic use of KPI data in budgeting and investment decisions	Similar to IBNET/CWIS, completed strategic sanitation planning exercises and feasibility studies/investment plans; decision-making process for sanitation budget allocation is transparent, inclusive, and informed by city/service area strategies; quality of investment decision-making	IBNET/CWIS	Related to political economy	Generally subjective measures, might be resource-intensive to assess well	Low
Finance, Transport	Average sludge/septage tipping fee at treatment site	Average cost per volume of fecal sludge delivered to treatment plant; alternatively, profit margin for private emptiers	CWIS/IBNET	Easy to track	Not as relevant as cost-recovery to financial health, does not directly measure profitability for private emptiers	Low
Finance	Operating cost recovery	Percentage of operational costs recovered through customer tariffs and revenue-generating activities (or profit margin, in Bangladesh) for FSM overall or treatment only (CWIS); should consider in combination with government cross-subsidies (e.g., from water/sewer bills, taxes)	CWIS; ESAWAS 2022a; Zambia "Monitoring & Performance Reporting," n.d.; SNV Bangladesh 2021a; Key informants (Philippines, Haiti)	Related to long-term financial sustainability	May differ among locations due to customer make-up	High
Finance	Unit capital and recurrent costs (per person) for construction and maintenance per year for all sanitation systems	From CWIS: Percentage of sanitation capital investments covered by budget line or government transfers	African Sanitation Policy Guidelines (Annex 1.2); CWIS	Indicates long-term financial planning	Hard to normalize over time; depends on construction cycles	Medium
Finance	Total annual investment in on-site sanitation	Budget amount allocated for on-site sanitation service provision	African Sanitation Policy Guidelines (Annex 1.2)	Potentially helpful for advocacy (e.g., to finance ministers), may align with Global Analysis and Assessment of Sanitation and Drinking-Water (GLAAS) report	Hard to understand relative value and adequacy of investments at different scales	Medium

Categories	KPI	Definition or Measurement Guidance	Source	Pros	Cons	Priority*
Finance	Bill collection efficiency	Percentage of revenue received divided by revenue owed for services	ESAWAS 2022a; Zambia "Monitoring & Performance Reporting" n.d.	Aids cost-recovery and financial sustainability	Cash transactions may be harder to track accurately and dependent on billing staff capacity, locations, and procedures	Medium
Finance	Budget line item for on-site sanitation	Local service authority on-site sanitation budget is a separate line item independent of water, solid waste management, health, or environment; local authority's sanitation revenue is ringfenced	CWIS (adapting to on-site sanitation and city- versus national-level indicators)	Critical to financial sustainability	Setups may differ among contexts	Medium
Finance	Staff efficiency	Average personnel cost per number of staff per month; New IBNET, ESAWAS, and NWASCO track the number of employees per 1,000 customer households	New IBNET; ESAWAS 2022a; Zambia "Monitoring & Performance Reporting" n.d.	May address one category of expenditures	Highly variable among contexts depending on level of automation, labor laws, etc.	Low
Finance	Percentage of clients who pay on time	End-of-year percentage of paid accounts divided by all accounts	Key informants (Haiti)	Indirectly related to affordability and service satisfaction	Amount of time allowed for account payment is likely to differ depending on predominant local industries (i.e., when customers gain income during the year)	Low
Finance	Timely payments	Percentage of payments for FSM service (e.g., truck contractors, treatment plants) issued by the local government within 3 months	Velidandla et al. 2020	Incentivizes private sector participation	May not be directly controlled by FSM unit of governance	Low
Finance	Budget utilization	Percentage of on-site sanitation budget spent (adapted from WASH indicator)	Key informants (Kenya; "RUSH - Rural Urban Sanitation and Hygiene" n.d.)	Reflects service provider level of effort	May not directly align with efficiency and depends strongly on adequacy of total budget	Low
Finance	Liquidity ratio	Assets divided by liabilities based on average from last four quarters	SNV Bangladesh 2021a	Measure of financial health	Debt may be inherited or related to larger economic crises	Low
Management, Collection	Emptier compliance with standard operating procedures	Number of mechanical pit latrine emptiers that provide services according to standard operating procedures	African Sanitation Policy Guidelines (Box 27)	Ensures consistently safe practices for workers and customers	Would require a skilled auditor	Low
Management, Finance	Performance agreements in place	Number of service providers with performance agreements in place with the government	African Sanitation Policy Guidelines (Box 27); CWIS/IBNET	Some promising examples of performance	Performance-based contracts do not work in all scenarios	Low

Categories	KPI	Definition or Measurement Guidance	Source	Pros	Cons	Priority*
	with service providers			agreements motivating service improvement	(may be demotivating if not well-targeted)	
Management	Performance indicators monitored and reported	Service authorities actively report on the performance indicators; public access to service provider performance data	CWIS; key informants (Philippines)	Measures upward accountability for benchmarking	Does not necessarily create clarity around how the regulator uses the information or provides feedback to the service provider	Medium
Management	Performance data from service authorities are transparent	Service authorities regularly collect representative data	CWIS	Bolsters public accountability and data reuse (e.g., for research)	Local laws may limit sharing	Medium
Management	Percentage of licensed service providers	Number of licensed providers divided by total number of estimated providers (all sanitation service providers, including those involved in constructing toilet facilities, emptying, transport, and treatment services, are operating under a sanitation service license); alternatively, from IBNET: Regulation of local private sector service providers (e.g., emptiers) via contractual arrangements or other regulatory mechanism	IBNET; African Sanitation Policy Guidelines (Box 27)	Reflective of professionalization, likely related to customer satisfaction	Hard to accurately estimate number of unlicensed providers	Medium
Management	Extent to which the provision of sufficient, safe, acceptable, accessible, and affordable WASH services are prioritized in national plans and policies	Presence of national FSM policy and guidance; similar to multiple CWIS indicators (e.g., policy, strategy, regulations for on-site sanitation developed and/or operationalized; required functions (operating/financial powers) delegated from government to sanitation authority/service provider; legal mandate for service delivery is clear and inclusive; mandated service authorities are delivering inclusive services; clear financing framework at the national level; integrated CWIS strategy	African Sanitation Policy Guidelines; IBNET/CWIS	Demonstrates intent for nationwide coverage	Presence/absence might provide shallow information; spectrum of national prioritization would be hard to define and is constantly shifting	Low
Management	Regulatory enforcement of	Incentives and/or penalties are actively used/applied by national/state	CWIS (adapting to on-site sanitation and city-	Enforcement activities are likely to drive	Hard to monitor quantitatively; may be difficult for regulator to act	Low

Categories	KPI	Definition or Measurement Guidance	Source	Pros	Cons	Priority*
	incentives or penalties	accountability authorities at the service area level	versus national-level indicators)	policy compliance in practice	comprehensively given limited staff capacity	
Management	Percentage of KPI performance targets met	Clear and progressive performance targets are set, monitored, and enforced for on-site sanitation	CWIS (adapting to on-site sanitation and city- versus national-level indicators)	Measures organizational accountability for benchmarking	Performance largely depends on how ambitious the targets are	Low
Management	Adequate staffing	Approved local service authority staff positions within mandated authority areas are sufficient to execute regulatory mandate to provide on-site sanitation services; local service authority staff positions are filled and capable to execute mandate	CWIS	Might identify a specific barrier to service delivery	Subjective; need a validated measurement scale for each context	Low
Environmental	Greenhouse gas emissions from desludging trucks (transport) and treatment	Estimated carbon emissions from vehicles per unit traveled times distance traveled plus estimated emissions from on-site sanitation storage types multiplied by the number of systems and residence time plus emissions from the treatment plant	Manila Water 2022, Urban Resilience by Building and Applying New Evidence in Water, Sanitation, and Hygiene (URBAN WASH) Activity	Keeps climate goals visible; may help to optimize transit time for desludging operators	Requires estimated emissions and data extrapolation; difficult to measure accurately	Medium
Environmental	Infrastructure vulnerability to climate change	Percentage of on-site sanitation facilities vulnerable to flooding	Staff exercise	Could be addressed within construction standards (codes)	Unclear how to measure; may require on-site visits	Low
Environmental	Greenhouse gas reduction from briquette sales	Estimated carbon emission reduction per unit multiplied by number of units sold	Staff exercise	Potential ties to climate finance opportunities	Proxy estimate; actual usage of products may vary; location-specific	Low

*Priority scale: High = used in multiple country programs or integral to global monitoring efforts, represents a critical FSM service aspect, and has few measurement drawbacks; Medium = used in one or more country programs, with a relatively even split between pros and cons; Low = limited use in practice, not critical to FSM service, or has limitations that would challenge consistent use across contexts.

U.S. Agency for International Development
1300 Pennsylvania Avenue, NW
Washington, DC 20523
Tel: (202) 712-0000
Fax: (202) 216-3524
www.usaid.gov