PLOS WATER

RESEARCH ARTICLE

Household, neighbourhood and service provider risk factors for piped drinking-water intermittency in urban and peri-urban Zambia: A cross-sectional analysis

Mair L. H. Thomas-Possee[1,2]*, Andrew A. Channon[3☯], Robert E. S. Bain[4], James A. Wright[1☯]

1 Geography and Environmental Science, University of Southampton, Southampton, United Kingdom, **2** WaterAid, London, United Kingdom, **3** Social Statistics and Demography, University of Southampton, Southampton, United Kingdom, **4** UNICEF Middle East and North Africa, Amman, Jordan

☯ These authors contributed equally to this work.
* mairthomas-poseee@wateraid.org

Citation: Thomas-Possee MLH, Channon AA, Bain RES, Wright JA (2024) Household, neighbourhood and service provider risk factors for piped drinking-water intermittency in urban and peri-urban Zambia: A cross-sectional analysis. PLOS Water 3(2): e0000127. https://doi.org/10.1371/journal.pwat.0000127

Editor: Sara Marks, Eawag, SWITZERLAND

Received: March 23, 2023

Accepted: October 4, 2023

Published: February 5, 2024

Data Availability Statement: The original contributions presented in the study are publicly available. The data can be found at the following: Service provider data can be accessed at: https://database.ib-net.org/countries_results?ctry=29&years=2018&type=report&ent=country&mult=true&report=1&table=true&chart=false&chartType=column&lang=EN&exch=1. 2018 DHS data is available on application via the public repository: https://dhsprogram.com/methodology/survey/survey-display-542.cfm. Regulator data is openly available and can be extracted from the

Abstract

Given nearly one third of sub-Saharan Africa's population lack access to an improved water source that is available when needed, service continuity restricts access to safely managed services. Household surveys, water regulators, and utilities all gather data on service continuity, but few studies have integrated these disparate datasets to quantify continuity-related risk factors and inequalities. This study aimed to assess the added value of utility and regulator data for international monitoring by assessing factors affecting piped water availability in urban and peri-urban Zambia. Household 'user' data from the 2018 Demographic and Health Survey (n = 3047) were spatially linked to provider data from an international utility database and regulator reports. Multilevel modelling quantified provider-related and socio-economic risk factors for households reporting water being unavailable for at least one day in the previous fortnight. 47% (95% CI: 45%, 49%) of urban and peri-urban households reported water being unavailable for at least one full day, ranging from 18% (95% CI: 14%, 23%) to 76% (95% CI: 70%, 81%) across providers. Controlling for provider, home ownership (odds ratio (OR) = 1.31; p <0.01), speaking Luvale, Kaonde, Lunda (OR = 2.06; p <0.05) or Tonga (OR = 1.78; p <0.1) as an ethnicity proxy, and dry season interview dates (OR = 1.91; p <0.05) were associated with household-reported interruptions. Households using a neighbour's tap (OR = 1.33; p <0.1) and in mid-wealth neighbourhoods (OR = 4.31; p <0.1) were more likely to report interruptions. For every $1000 increase in utility-level GDP per capita, the odds of an interruption were 0.51 times less (p<0.01). Substantial inequalities in drinking-water availability were found between provider coverage areas. Spatial integration of user, provider and regulator data enriches analysis, providing a finer-scale perspective than otherwise possible. However, wider use of utility or regulator data requires investment in monitoring of small-scale community supply intermittency and utility coverage area data.

following report: https://www.nwasco.org.zm/index.php/media-center/publications/water-supply-and-sanitation-sector-reports. All code and publicly available data has been provided as supplementary materials to allow full transparency and replication of the analysis that has been undertaken.

Funding: This research was supported by a South Coast Doctoral Training Partnership (SCDTP) PhD award, awarded to M.L.H.T-P., from the UK Economic and Social Research Council (Grant No. ES/P000673/1). This funding body had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Introduction

The sixth Sustainable Development Goal (SDG) aims to 'ensure the availability and sustainable management of water and sanitation for all' by 2030 [1]. The accompanying target 6.1, is to achieve universal and equitable access to safe and affordable drinking-water. The associated indicator, 6.1.1, measures progress towards this target via the proportion of the population using safely managed drinking-water services, requiring that drinking-water from an improved service be available when needed [2]. Additionally, the World Health Organization (WHO) outlines a global drinking-water availability benchmark which recommends that a minimum of 50 litres/capita/day (LPCD) is needed to meet domestic needs, including washing, personal hygiene and cleaning [3, 4].

Between 2000 and 2020 the proportion of the sub-Saharan African (SSA) population using improved drinking-water sources that were available when needed increased from 41% to 59% [5]. Over the same period, the population in SSA using piped water doubled from 185 to 380 million. However, coverage of piped water services in urban areas has not kept up with population growth, declining from 65% to 59% between 2000 and 2020 [5]. In 2015, an estimated 116 million people in Africa were supplied by an unreliable piped water system prone to interruptions [6]. Water services that are not available when needed may result in damage to water service infrastructure, compromise water safety [7], adversely impact health [8, 9], and lead to additional household expenditure on water storage, treatment and purchasing supplementary water [10], with the latter often sourced from informal service providers and unimproved services [11].

The transition from the Millennium Development Goals (MDGs) to the SDGs saw the specific addition of water availability to the international agenda, resulting in new demand for data sources for monitoring [12]. Multiple sources, including household surveys (hereafter 'user data'), utility companies (hereafter 'provider data') and government regulators [13], hold data on water service intermittency. Historically household survey questions on the availability of drinking water services have not been harmonised, thus different countries may not use consistent question and response wording. This complex data landscape is further exacerbated by the challenges of measuring water availability, for which methods vary between studies [14]. Metrics used include hours of service a day alongside service in the last week or month, using household or per capita consumption per day, or the number of interruptions or breakages in a given time period [15]. In 2018, the Joint Monitoring Programme (JMP) of the WHO and United Nations Children's Fund (UNICEF), who are responsible for international monitoring of SDG target 6.1, published a core question, '*In the last month, has there been any time when your household did not have sufficient quantities of drinking-water when needed*?', for greater harmonisation and incorporation into household surveys for availability monitoring. Since this is a new question and household surveys are implemented only every three to five years, availability of internationally comparable survey data remains patchy.

As a result of patchy data and the use of multiple metrics of availability, national and international monitoring has been reliant on numerous data sources, presenting challenges for data integration. Under the MDGs, monitoring was primarily dependent on household surveys and census data [16], whereas more recently under the SDGs, there has been a shift towards using information from regulators of providers alongside these more traditional sources [17]. Regulators often produce annual reports which benchmark levels of service between different providers. An additional data source concerning piped water service levels is available directly from service provider records [18]. Many service providers report their performance data to the International Benchmarking Network for Water and Sanitation Utilities (IBNET). At present, IBNET provides the most systematic international data on provider reported water availability [18].

To date, most studies that have sought to quantify risk factors for interruptions in water services have focused almost exclusively on household surveys [19–21]. In comparing provider and user data, one study conducted bespoke household surveys within four case study provider coverage areas (PCAs) in Kenya and Ghana [11], whilst a second compared user-reported and provider-reported service continuity at regional level in Peru [18]. However, since many household surveys are now georeferenced [22], the new household survey question on service continuity provides an opportunity for spatial integration of household reports of water service interruptions with related statistics reported by service providers and regulators. Such spatial integration would enable more detailed assessment of household versus regulator- or provider-reported service continuity for consistency, alongside evaluation of provider-level risk factors for household-reported interruptions such as supply-side service management indicators.

Barriers inhibiting household access to uninterrupted water services may exacerbate urban-rural and socio-economic disparities that are evident in drinking-water services. In 2020, only 13% of the rural SSA population used services that were safely managed compared to 54% of the urban population [5]. With the SDGs seeking to '*leave no one behind*', identifying and addressing drivers of inequalities in water, sanitation and hygiene (WASH) services is critical. In a city-scale study of Lilongwe, Malawi, water rationing of the formal water system led to more irregular supply in low-income informal settlements [23], highlighting one driver of such inequalities. However, whilst wealth quintile was associated with rural household access to continuous water in Bangladesh and Pakistan [19], there is little evidence on urban or per-urban inequalities in household access to continuous piped water in addition to such inequalities at city or PCA scale.

The primary objective of this study was therefore to assess the added value of utility and regulator data in estimating piped water availability and quantifying risk factors for piped water intermittency and related inequalities. As secondary objectives, the study aims to assess whether household survey data are sufficient in isolation to quantify water availability and related risk factors, alongside identifying barriers to more widespread integrated analyses of household survey, utility, and regulatory data sets. Data on water service availability from users, providers and regulators were first mapped to identify a suitable case study country to examine risk factors for user-reported interruptions in piped water services in SSA. Since Africa shows the least progress towards SDG 6 [5], and Zambia had near contemporaneous data from users, providers and regulators that could be spatially linked, it was selected as the case study.

Materials and methods

Ethical approval

Ethical approval for the use of all data was received from the University of Southampton Ethics Committee (Submission ID # 55516) on 10th March 2020.

Study country selection

A systematic secondary data audit (S1 Fig; S1–S3 Tables) was undertaken via the JMP's 2019 country files [24] and IBNET database [25] to identify an African Union (AU) member state for which household survey or population census data, government regulator data and water providers all report at least one water availability metric (i.e. a measure of the quantity of household water supplied, or continuity of water service provision) covering the same year(s). To facilitate subsequent multi-level analysis, countries eligible for inclusion required household survey or census data geographically disaggregated to sub-provincial level and provider or regulator data disaggregated to at least province level or equivalent coverage area.

Half of AU member states (n = 27) lacked relevant household survey or census data on availability and were excluded (Fig 1). Of the remaining 27 that had user data, only five had user, provider and regulator data that each captured water availability. For each of the five shortlisted countries, the three data streams were mapped using ArcMap 10.7.1. This

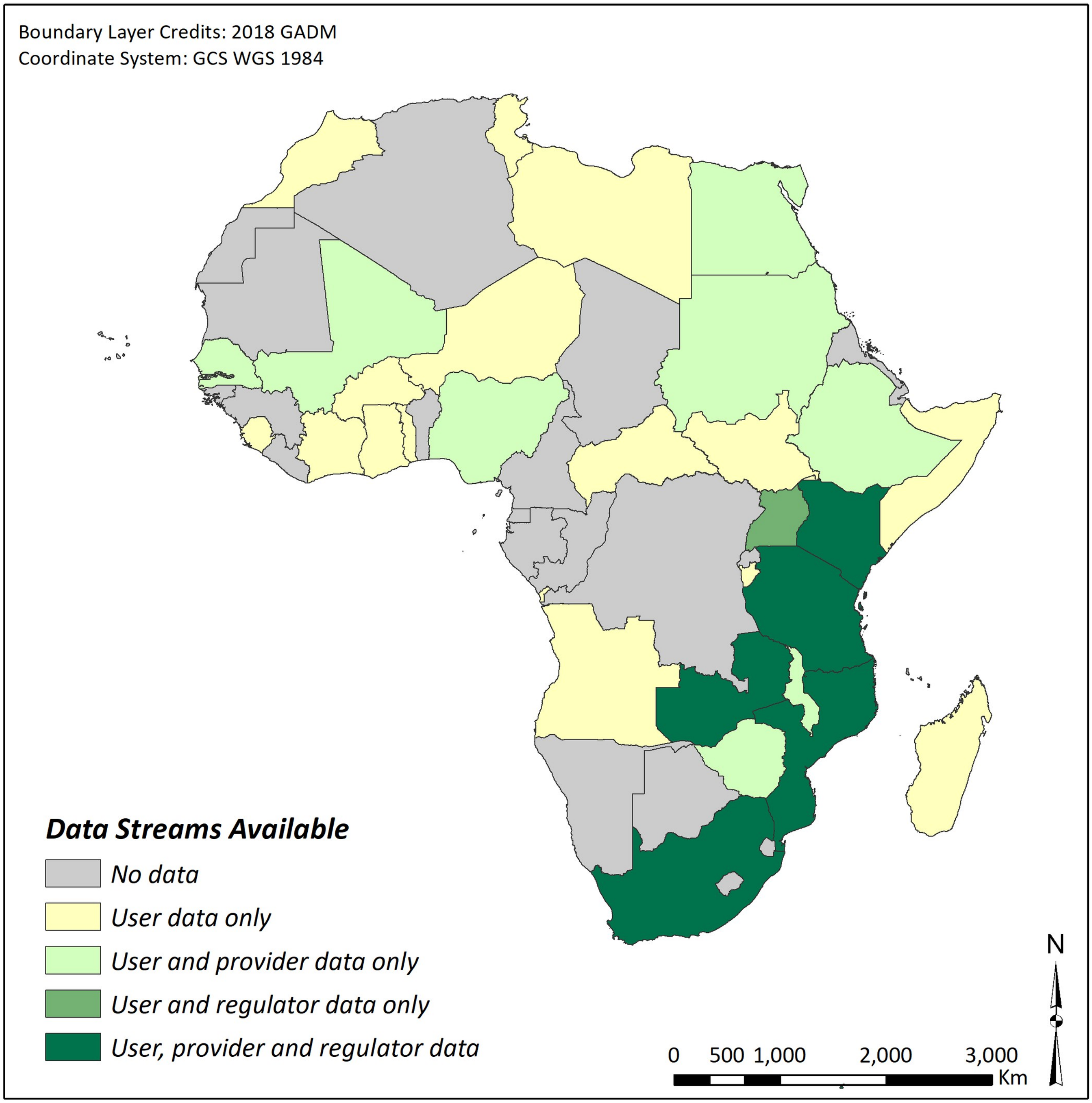

Fig 1. Available data for each of the 54 African Union member states. (GADM 2018 base layer available at: https://gadm.org/index.html).

https://doi.org/10.1371/journal.pwat.0000127.g001 determined whether subsequent integration of the three data streams was possible. Zambia was the only country where the administrative geography of drinking-water supply was sufficiently simple to facilitate spatial linkage of household survey and provider data since PCA (which are also used by the regulator) coincided with province boundaries. All three Zambian data sources also covered the same year (S1 Table).

Study site

Located in southern-central Africa, Zambia is a landlocked country with a 2019 population of nearly 19 million [26]; 45% live in urban areas [27]. Zambia's trans-boundary catchment areas result in fluctuations of surface water availability because of the varying water demands of neighbouring states [28]. Renewable water resources are affected by inconsistent seasonal rainfall, characterised by periodic drought [29]. Mismanagement and rapid urban growth have also caused considerable stress on groundwater resources [30].

Urban and peri-urban piped drinking-water is supplied to over six million people by 11 commercial Water and Sewerage Companies (WSCs) [31]. Approximately 4,656,375 piped water connections supply drinking-water in 91 towns (S4 Table). In total, an estimated 46% of the urban population have a safely managed water service, that is available when needed [32]. In rural Zambia, 35% have an unimproved service compared to 9% of the urban population [32]. The National Water Supply and Sanitation Council (NWASCO), a statutory body, is responsible for regulating water and sanitation services across Zambia [33].

Data sources and availability metrics

We used three main data sources in our study:

1. **User data:** nationally representative data from the Zambian Demographic and Health Survey (DHS) was used. This was implemented as a multi-stage cluster household survey from 18th July 2018 to 24th January 2019 [34]. Interviews were undertaken concurrently across all provinces, using a stratified two-stage sample design [34]. We used geospatial data, georeferenced to cluster level (comprising groups of approximately 25 households, selected at random from a given census enumeration area). Cluster locations are provided as the mean GPS coordinates for all participating households within each cluster, displaced within 2km (for urban areas) to retain anonymity [35]. No sampling clusters were displaced outside of their administrative district. To assess drinking-water availability, participating households are asked '*In the last two weeks, was the water from your main source not available for at least one full day*?'.

2. **Provider data:** To assess provider-level risk factors for reported interruptions, provider-reported availability of piped services for all 11 WSCs was obtained from the IBNET database for 2017, the most recent year available [36]. This includes metrics of availability, such as continuity of service, as yearly average hours of service per day (hrs/day) [37], and yearly average residential consumption in LPCD [38]. Details of non-revenue water, the difference in water supplied and sold as a percentage of net water supplied [39], which inherently affects service availability [40, 41] were also of interest as these reflect illegal use of piped networks and water that has been stolen or leaked from the system.

3. **Regulator data:** since 2001, NWASCO have also annually reported on each WSC's performance based on nine key indicators [33]. The annual average duration of water service (hours/day) at the customer connection, and the average amount of water consumed in LPCD were extracted from the publicly available 2018 NWASCO report [42] for each WSC.

Household inclusion criteria and data integration

With less than 5% of rural households using piped drinking-water services [5], provider and regulator data from NWASCO do not cover rural areas [43]. Given this, 8,117 households in 347 rural clusters (comprising 16% of households nationally reporting piped water as their drinking-water source) were excluded (Fig 2). 47 households in two clusters which had missing GPS coordinates were also excluded. Provider service area boundaries for regulator and

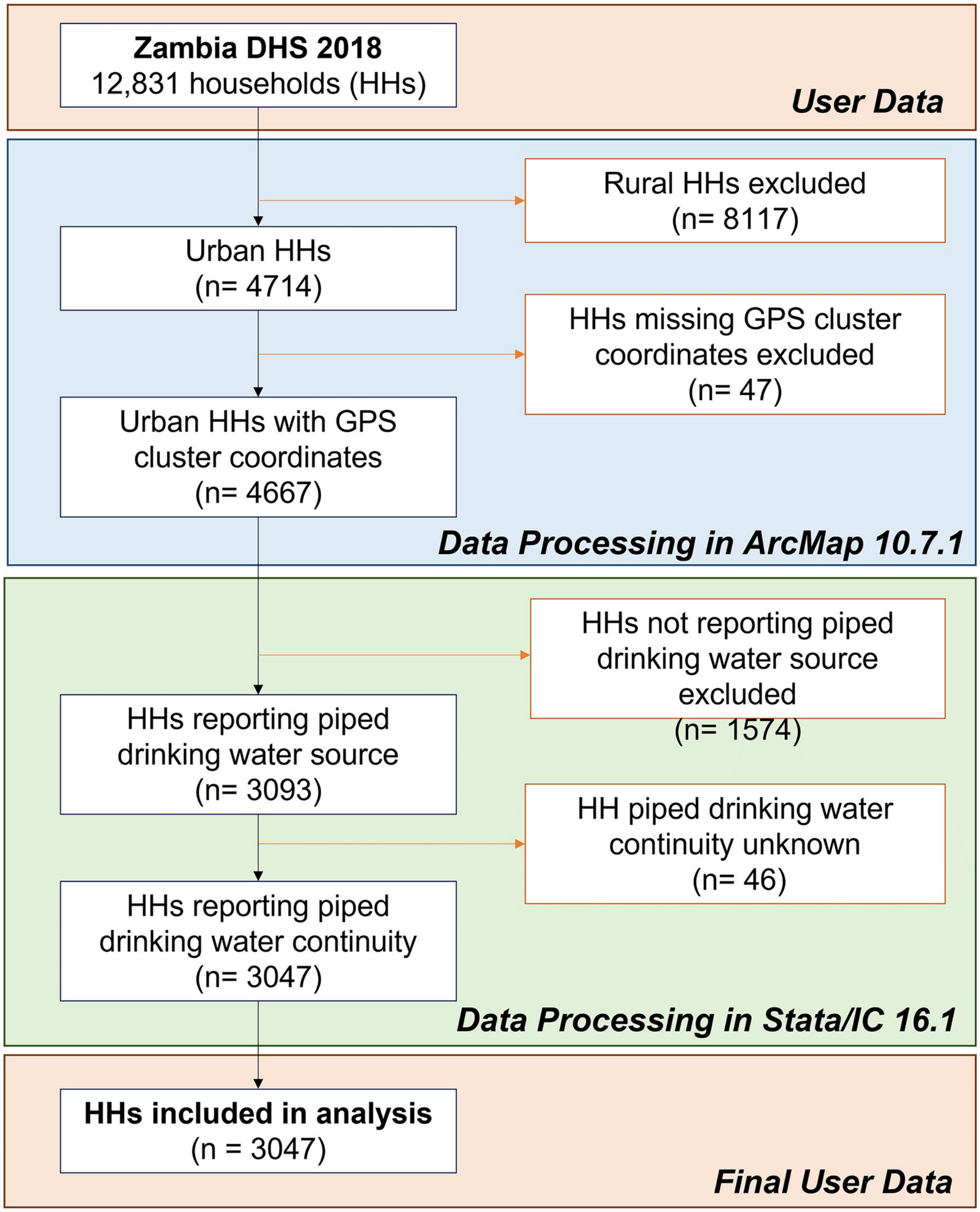

Fig 2. Flowchart, showing reasons for excluding households participating in the 2018 Zambian Demographic and Health Survey from analysis (n = number of households).

https://doi.org/10.1371/journal.pwat.0000127.g002

provider data were derived through aggregation of district administrative boundaries for Zambia from the Global Administrative Areas (GADM) database [44] and spatially joined to DHS household clusters in ArcMap 10.7.1. The final dataset therefore included information on the provider and regulator perspective of water availability at the PCA level (henceforth 'provider-level') and the user perspective at household-level.

Households that did not drink piped water (whether from a connection in dwelling, yard/plot, neighbour or public tap/standpipe) were excluded (n = 1574) as they were not asked about water service continuity [45], as were those who did not know about their piped water continuity (n = 46) (Fig 2).

Outcome and explanatory variables

A binary outcome variable 'user reported availability of piped water' was defined as respondents reporting at least one full day of interruptions in the two weeks prior to being surveyed. Explanatory variables at the household-, cluster- and provider-level were chosen to represent socio-economic, water source, or neighbourhood characteristics that could constitute risk factors for piped water interruptions (S5 and S6 Tables). Household cluster-level variables were created by aggregating data on households within each DHS cluster.

In modelling user-reported availability of piped water, the following household-level explanatory variables were included:

- Since different types of connection are prone to different types of interruptions [46], ***type of piped service*** (piped into dwelling, piped to neighbour, piped to yard/plot or public tap/standpipe) was included as a covariate. Users may also consume more water if it is piped to the yard or home, than if located further afield, such as at a neighbour's or a public tap/standpipe [47].
- The DHS's urban/rural wealth index was used as a ***household wealth*** variable. This considers assets and services owned by rural populations alongside those owned by urban populations [48]. Wealth relates to the type of piped service used [49] and affects household vulnerability to interruptions. The wealthier pay tariffs more regularly, are less exposed to illegal connections or pipeline breakages, can afford to consume more water [50] and purchase storage tanks which may protect from interruptions [51].
- ***Month of interview*** was included as a proxy for seasonal water shortages [52]. Lower borehole yields seasonally restrict groundwater-fed piped systems, while lower reservoir levels seasonally affect piped networks drawing on surface waters [53]. Date of interview was grouped into categories to align with the April-September dry season in 2018/19 [54] (dry season: 'September 2018'; 'July-August 2018', rainy season: 'October-November 2018'; 'December 2018-January 2019').
- Belonging to a minority ethnic group may restrict water access, since for example, locations of public standpipes and household connections to mains services often disproportionately favour majority ethnic groups [55]. ***Native language*** (English, Bemba, Lozi, Tonga, Kaonde/Lunda/Luvale or other) was therefore included as a proxy for ethnicity.
- ***Household size*** was included as it correlates with water consumption [51, 56], including depletion rates of household stored water, ultimately affecting continuity. It also relates to choice of service type [57, 58].
- ***Home ownership*** relates to the choice of water service and ability to cope with piped water interruptions. For example, home-owners can choose to invest in a water tank as an

important coping strategy [59]. Home ownership as reported by the men's and women's DHS questionnaire respondents was included.

At DHS household cluster level, we included the following variables:

- ***Degree of urbanisation*** (classified as: urban centres/cities; urban clusters/towns/suburbs; rural localities; and unpopulated) was included as a systematic review found lower reported availability in rural settings [15]. This DHS variable was derived from 2015 settlement data from the Global Human Settlement Layer (GHSL) project [60].
- Since recently urbanised areas may have water infrastructure that has undergone development and is less prone to maintenance-related interruptions, we included ***change in urbanisation*** (less urbanised; more urbanised; no change), based on the difference in GHSL settlement class between 1990 and 2015.
- ***Neighbourhood wealth*** was included to reflect societal factors such as crime levels and disadvantaged communities [61], which in turn affect neighbourhood water infrastructure and local capacity to cope with water intermittency. Cluster-level averages of the DHS's urban/rural wealth index were created.
- Newer neighbourhoods are less likely to have water infrastructure that is prone to failure, due to general aging of materials and poor upkeep [11, 62] and neighbourhoods comprising newly arrived migrants may lack sufficient social cohesion to lobby for services [63]. The men's and women's DHS questionnaires ask respondents their length of residence in their current home. As a proxy for the age of a cluster's neighbourhood and social cohesion, we calculated the maximum ***length of residency*** for any men or women within a given cluster.

Water service provider-level variables were as follows:

- Since local economic development affects investment in WASH infrastructure [46, 64], we calculated ***GDP per capita at the service provider level***. 2015 provincial GDP [65] data were converted to GDP per capita using 2015 provincial population statistics [66] (S7 Table), estimating this via areal interpolation for those service provider areas that did not match to provincial boundaries.
- ***Non-revenue water*** was included to account for leakage or illegal use of piped networks [40, 41]. 2017 IBNET data which comprised the difference between water supplied and water sold that is 'lost' before it reaches the consumer, expressed as a percentage of net water supplied [39].
- ***Provider and regulator availability***–reported average service hours expressed as hrs/day and LCPCD were used and treated as continuous variables.

Statistical analysis

Stata 16.1 was used for all analysis [67]. Initially, provider, regulator and user-reported availability were compared through scatter plots and calculation of Pearson's correlation coefficients. The LPCD measures reported by providers and the regulator were also assessed against the WHO benchmark of ≥50 LPCD to estimate the proportion of population within each PCA not meeting this benchmark.

Descriptive and bivariate analysis for the outcome and explanatory variables, and associated 95% confidence intervals (CIs) and chi-square tests, used household survey weights [68] and accounted for the complex survey design. Collinearity and missingness were examined for the explanatory variables.

Comparison of a single-level unconditional model and a two-level variance components model quantified the clustering in the dataset, confirming the appropriateness of multilevel models. Households (level 1) were nested in household clusters (level 2). Two final models were specified. Model 1 included significant household and household cluster level variables, and accounted for provider-level variation using dummy variables, as the small number of PCAs (<25) made it inappropriate to include the PCA as a level [69]. Model 2 included significant household, household cluster and provider explanatory variables. Backward elimination with a significance threshold of p = 0.05 was used for both models and type of piped service was retained as a control, despite not reaching this level of significance. An interaction term was considered when a valid hypothesis existed and was subsequently deemed meaningful. Survey weights were not used in the multilevel models, since the DHS only make weights available at the household level due to concerns about disclosure risk [70]. Therefore, they were unavailable at the household cluster or provider level, as was required for this analysis. Lastly, random effects were added to the two models to calculate the proportion of variance within household clusters that could not be attributed to observed variables.

Issues of multicollinearity were found between regulator- and provider-reported LPCD or service hours/day. Analysis found regulator and provider reporting for the same year (2017) to be very similar for LPCD, and identical for hrs/day (see section entitled *'Data Sources and Availability Metrics'*). Detailed analysis found a clear relationship between regulator LPCD in 2017 and 2018 (S2 Fig). LPCD was therefore used as the measure of availability. When comparing regulator and provider perspectives during initial descriptive analysis, 2017 data was used for consistencies of the year. For all modelling however, to eliminate multicollinearity, only 2018 regulator data was used as it matched the year of user-reported availability and reflected the provider perspective.

Results

Comparison of provider, regulator and user availability

In 2017, regulator records for service continuity (hrs/day) perfectly matched the provider data reported to IBNET. No utilities reported a continuous service for 24 hrs/day, with households receiving piped water for 18.4 hrs/day on average. The regulator and providers reported similar quantities of water supplied (in LPCD). Reported quantities were positively correlated (r_s = 0.86, $p<0.001$), though higher average LPCD was reported by providers (S16 Table).

2018 regulator-reported service continuity (hrs/day) is associated with user-reported availability at PCA level ($\chi(4) = 56.04$, $p<0.001$). Where the regulator reported continuity of 20 hrs/day, 54% of users reported interruptions. 2018 regulator LPCD and user-reported availability were associated ($\chi(9) = 220.63$, $p<0.001$), with evidence of a significant weak positive correlation (r_s = 0.05, p = 0.007). Where the regulator reported supplying ≥50 LPCD and subsequently met the WHO benchmark, 55% of users reported an interruption to their supply.

User-reported availability by provider coverage areas

Overall, 47% (95% CI: 44%, 49%) of users reported experiencing at least one full day of interruptions in the two weeks prior to being surveyed in 2018. The proportion of households reporting an interruption ranged from 77% (95% CI: 71%, 82%) in Kafubu WSC to 19% (95% CI: 15%, 24%) in Lukanga WSC (S8 Table). Fig 3 shows 2018 user-reported availability alongside regulator/provider-reported hrs/day. Only six of 177 household clusters had no household reports of piped water interruptions. Lusaka WSC supplied piped drinking-water to 44% of households in the sample. Southern WSC supplied a further 14% of sampled households. All other WSCs provided drinking-water to between 2% and 9% of sampled households.

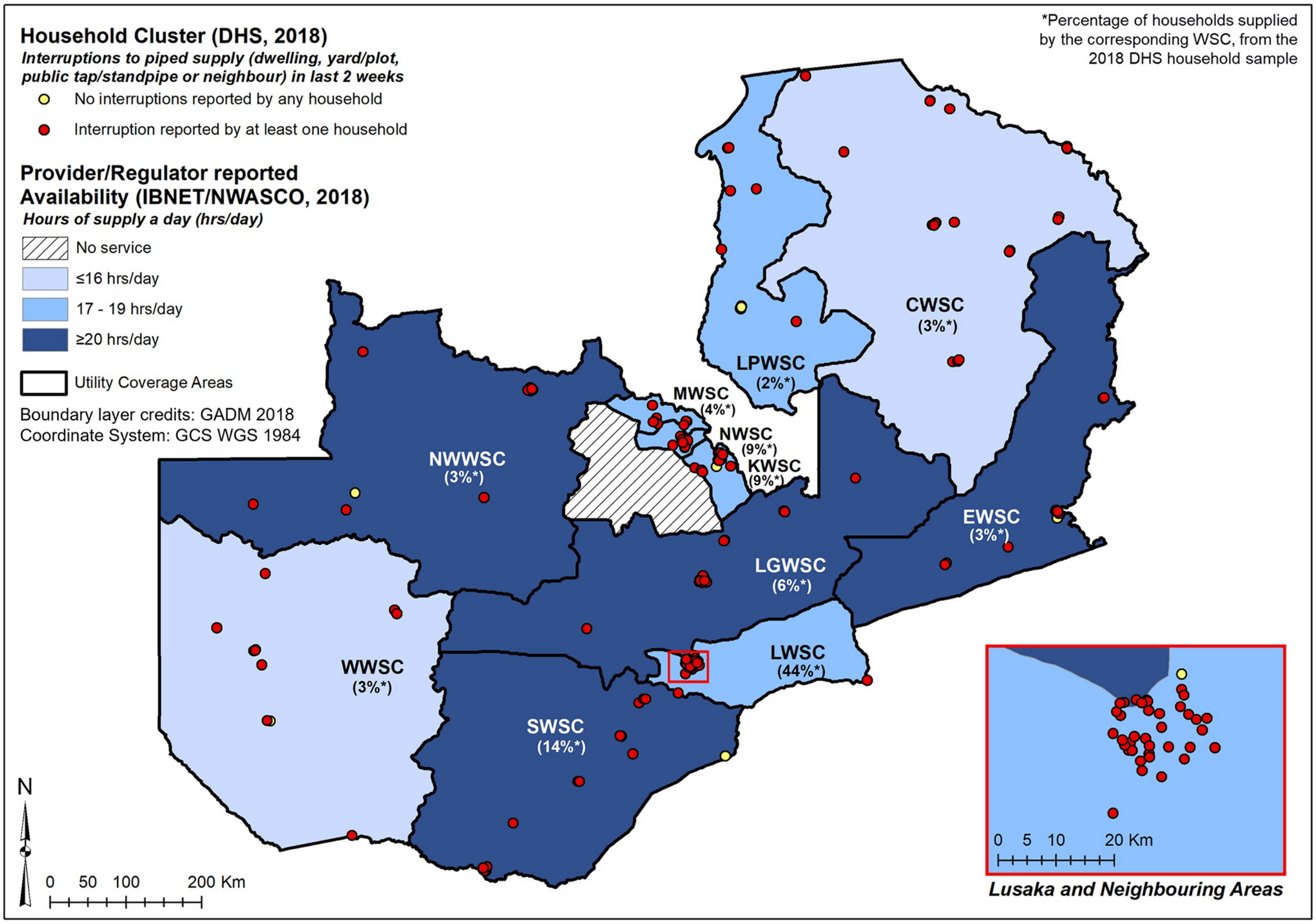

Fig 3. Household and regulator/provider-reported availability of piped water in Zambia (2018). (Key to Utilities/Water and Sewerage Companies (WSC): NWWSC -North Western WSC; MWSC- Mulonga WSC; NWSC- Nkana WSC; KWSC- Kafubu WSC; LGWSC- Lukanga WSC; LPWSC- Luapula WSC; CWSC- Chambeshi WSC; ESWC- Eastern WSC; LWSC- Lusaka WSC; SWSC- Southern WSC; WWSC- Western WSC) (GADM 2018 base layer available at: https://gadm.org/index.html).

https://doi.org/10.1371/journal.pwat.0000127.g003

Bivariate analysis of user-reported water availability

1398 of 3047 households included in analysis reported an interruption to their service, with the proportions of such households varying seasonally and by respondent's native language (Table 1).

Neither household wealth, household size, home ownership nor type of piped service were statistically significantly related to service interruptions ($p>0.05$). Month of interview was significantly associated with interruptions to drinking-water service ($p<0.05$), as was native language ($p<0.05$). As expected, type of piped service was significantly associated with household wealth ($p<0.01$). 68% of households in the richest wealth quintile had a piped service into their dwelling, compared to 2% in the poorest wealth quintile. 45% of those in the poorest wealth quintile used a public standpipe/tap compared to less than 1% in the richest wealth quintile.

At the household cluster level, only degree of urbanisation and maximum length of residency were significantly associated with interruptions ($p<0.01$) (Table 2). The relationship was non-linear, with reported water interruptions lowest at 32.2% for clusters with a

Table 1. Proportion of urban Zambian households reporting a piped water service interruption in the preceding fortnight, by socio-economic characteristic (n = 3047).

User Characteristics	Weighted Percentages (%) of Households (n) Reporting an Interruption	Chi squared (df), p-value
Household Size		3.54 (4), 0.47
1 person	42.6 (108)	
2–3 people	48.2 (358)	
4–6 people	47.2 (618)	
7–9 people	46.7 (250)	
10+ people	40.3 (64)	
Native Language		52.69 (8), <0.001**
English	25.8 (21)	
Bemba	50.3 (597)	
Lozi	49.1 (117)	
Lunda, Kaonde, Luvale	50.8 (137)	
Tonga	44.6 (142)	
Other	44.1 (384)	
Household Wealth		0.84 (4), 0.93
Poorest	45.5 (236)	
Poorer	44.3 (244)	
Middle	47.0 (273)	
Richer	48.8 (298)	
Richest	47.1 (347)	
Home Ownership		6.19 (2), 0.05
At least partly owns house	50.8 (477)	
Does not own	45.0 (855)	
No information	43.0 (66)	
Type of piped service		1.40 (2), 0.50
Piped into dwelling	47.7 (322)	
Piped to yard/plot or Public tap/ standpipe	47.0 (793)	
Piped to neighbour	44.7 (283)	
Month of Interview (Season)		17.56 (3), 0.001**
July-August 2018 (Dry)	49.8 (653)	
September 2018 (Dry)	43.8 (275)	
October-November 2018 (Rainy)	47.2 (349)	
December 2018- January 2019 (Rainy)	32.6 (121)	
Provider		245.87 (10), 0.93
Chambeshi WSC	57.3 (148)	
Eastern WSC	38.1 (78)	
Kafubu WSC	77.0 (166)	
Luapula WSC	58.2 (59)	
Lukanga WSC	18.7 (52)	
Lusaka WSC	42.4 (377)	
Mulongo WSC	58.8 (73)	
Nkana WSC	53.3 (121)	
North Western WSC	48.2 (105)	
Southern WSC	41.9 (135)	
Western WSC	42.7 (84)	

(*Continued*)

Table 1. (Continued)

User Characteristics	Weighted Percentages (%) of Households (n) Reporting an Interruption	Chi squared (df), p-value
Total households reporting an interruption:	46.7 (1398)	

Note: n is an unweighted count

* p<0.05

**p<0.01

https://doi.org/10.1371/journal.pwat.0000127.t001

maximum length of residency of 31–40 years. At the provider level, GDP per capita, regulator- and provider-reported service continuity, regulator- and provider-reported LPCD and non-revenue water were all associated with service interruptions ($p<0.01$).

User-reported inequalities in water availability by provider

Table 3 presents the socio-economic differences in household characteristics of those reporting interruptions to their service, by provider. Ratios of those reporting an interruption are used as a measure of inequality within each PCA.

Inequalities in those experiencing an interruption between providers are sometimes large, though not necessarily significant, when considering a range of household characteristics. Between PCAs, inconsistencies often existed in which household group reported the most interruptions. In 82% of PCAs more one-person households experienced interruptions than households with >10 members.

In all PCAs, more households using a public standpipe or a service in their yard experienced interruptions than those with a service in their dwelling. In Luapula WSC for example, households using a public standpipe or a service in their yard were 8.9 times more likely to experience an interruption. Households that at least partly owned their home were more likely to report an interruption than those that did not, in all but one PCA. Variation exists in the

Table 2. Association between household-reported availability and household cluster level and provider level explanatory variables in urban Zambia (n = 3047).

Explanatory variable	Chi-squared (df)	p-value
Household Cluster Level Factors		
Neighbourhood wealth	1.4 (2)	0.50
Maximum length of residency	53.9 (3)	<0.001**
Change in urbanisation	1.4 (2)	0.49
Degree of urbanisation	15.6 (3)	<0.001**
Water service provider-level Factors		
GDP per capita at service provider level	245.9 (10)	<0.001**
Regulator hrs/day	56.0 (4)	<0.001**
Provider hrs/day	144.3 (6)	<0.001**
Regulator LPCD	220.6 (9)	<0.001**
Provider LPCD	245.9 10)	<0.001**
Non-Revenue Water	214.7 (10)	<0.001**

Note

* p<0.05

**p<0.01

https://doi.org/10.1371/journal.pwat.0000127.t002

Table 3. User-reported inequalities in piped water interruptions, by provider.

Provider	Percentage (%) (weighted) difference of households reporting an interruption between population sub-groups with differing socio-economic characteristics						Ratio of households reporting an interruption between population sub-groups with differing socio-economic characteristics					
	Partly Owns Home versus Does not Own	Richest versus Poorest HHs	One person versus >10 people HHs	Yard/ Public Standpipe versus Supply in Dwelling	Interviewed in Dry (July/ Aug) versus Rainy Season (Dec/Jan)	Majority versus Minority Language in WSC Area	Partly Owns Home / Does not Own	Richest / Poorest HHs	One person / >10 people HHs	Yard/ Public Standpipe / Supply in Dwelling	Interviewed in Dry (July/ Aug) / Rainy Season (Dec/ Jan)	Majority / Minority Language in WSC Area
Chambeshi WSC	4.8	-29.5	0.8	21.4	16.5	93.0	1.2	0.3	1.1	1.6	1.9	65.8
Eastern WSC	20.6	-2.5	7.0	25.4	7.6	84.5	1.8	0.9	2.4	1.8	1.4	28.8
Kafubu WSC	32.3	20.8	3.7	19.2	45.2	63.1	2.6	2.9	2.1	1.9	6.3	45.4
Luapula WSC	15.7	-7.5	-5.4	62.6	32.6	84.4	1.6	0.8	0.2	8.9	4.5	21.0
Lukanga WSC	19.5	3.4	7.7	49.0	53.8	41.9	1.8	1.1	2.0	5.1	4.3	14.1
Lusaka WSC	39.0	21.0	3.7	44.9	48.8*	46.5	3.4	5.0	2.1	3.7	-*	116.1
Mulongo WSC	39.2	19.3	-1.0	55.9	29.8	73.7	3.3	4.1	0.8	7.0	8.8	69.4
Nkana WSC	44.0	15.4	1.0	55.9	55.6*	62.8*	4.0	3.4	1.2	5.8	-*	-*
North Western WSC	-0.1	17.9	6.5	40.1	52.0	53.5	1.0	2.3	3.7	3.3	10.1	44.4
Southern WSC	18.0	-19.5	10.7	16.2	63.4	47.6*	1.7	0.2	4.5	1.5	9.9	-*
Western WSC	21.6	-31.1	1.3	12.0	29.6	71.8*	2.4	0.4	1.3	1.3	2.2	-*

Negative difference values and ratios below 1.0 indicate lower reported interruptions in the first population sub-group named in column heading. Counts and weighted percentages of the number of households reporting an interruption for each inequality are presented in S9–S14 Tables. S15 Table & S3 Fig present household cluster inequalities in water availability by provider.

*Household did not report an interruption in one of the socio-economic categories being compared.

https://doi.org/10.1371/journal.pwat.0000127.t003

magnitude of this inequality (Table 3). In 55% of PCAs, the richest households experienced interruptions more than the poorest. Differences in household interruptions between those speaking the majority and minority language are consistent between each PCA. Households interviewed in the drier months of July-August were more likely to report an interruption than those interviewed in wetter December-January, regardless of PCA.

Some patterns of socioeconomic inequalities in households reporting an interruption may be mediated by PCA characteristics. In urbanised PCAs, the gap between households speaking the majority language that report an interruption, compared to the minority language, is narrower than for less urbanised PCAs. For both home ownership and household wealth, the richest PCAs had the greatest inequalities in service interruption risk. Conversely, when comparing households using a yard or public tap versus those with water piped to their dwelling, or those speaking the majority versus minority language, the richest PCAs had the smallest differences in those reporting an interruption.

Multi-level logistic regression

Table 4 includes the multi-level logistic regression results for two models. Model 1 includes selected household and household cluster factors that explain inequalities in piped-water interruptions with provider accounted for via dummy variables. Model 2 shows the association between socioeconomic characteristics at the household, household cluster and provider-level and the likelihood of households reporting an interruption.

When provider is accounted for in model 1, type of piped service is not significant, whereas in model 2 it is. Households using a neighbour's water had 33% greater odds of reporting an interruption than those with water piped to their dwelling ($p<0.1$). In model 1, households interviewed at the peak of the dry season had 1.91 times the odds of an interruption than those interviewed in the middle of the rainy season ($p<0.05$). In model 2, households interviewed outside the wettest months of December-January had 2–3 times the odds of experiencing an interruption ($p<0.05$).

In both models, household ownership is significantly associated with experiencing an interruption. The odds of experiencing an interruption were 1.31–1.32 times higher for those that at least partly owned their home ($p<0.001$). Availability of services varies significantly with household native language. In both models, those who speak Luvale/Kaonde/Lunda have greater odds of experiencing an interruption than English speaking households ($p<0.05$). When accounting for provider, compared to English speaking households, Tonga speakers were 1.78 times more likely to have an interruption ($p<0.1$) whereas in model 2, Bemba speaking households had 1.88 time the odds ($p<0.1$).

Households in neighbourhoods classed as having mid-level wealth had 4.31 times the odds of experiencing an interruption than those in poor neighbourhoods ($p<0.01$) (model 2). Neighbourhood wealth was found to interact with regulator-reported service availability. Compared to poor clusters, for every one litre increase in regulator reported LPCD, households in mid-wealth clusters had 3% lower odds of reporting an interruption ($p<0.05$).

GDP per capita is the only provider-level factor associated with household availability (model 2). For each $1000 increase in GDP per capita, the odds of reporting an interruption are 0.51 times less ($p<0.01$). GDP per capita at service provider level interacts with regulator-reported LPCD. For every one unit increase in regulator-reported LPCD and GDP per capita at service provider level, the odds of having an interruption were 1.02 times greater ($p<0.01$).

Substantial variations between providers are evident when controlling for home ownership, type of piped service, month of interview and native language (model 1). When comparing the odds of households reporting an interruption to areas supplied by Lusaka WSC, households supplied by Lukanga WSC had 64% lower odds of experiencing an interruption ($p<0.05$) whilst in Nkana the odds were 2.24 times higher ($p<0.05$) and Mulongo WSC 3.77 times higher ($p<0.05$). Similarly, in Chambeshi WSC the odds of having an interruption were 2.43 times higher ($p<0.05$) and in Luapula they were 2.89 ($p<0.05$). In Kafubu WSC, households have eightfold increased odds of reporting experiencing an interruption than those whose provision is from Lusaka WSC ($p<0.01$).

The variability in households' likelihood of experiencing an interruption that was not explained by the household and household cluster factors used in the multilevel models was examined through the Intra-Class Correlation (ICC). The ICC for model 1 shows that 27% of the remaining unexplained variation in reported interruptions lies between household clusters, whilst 73% lies within household clusters. In model 2, 31% of the variation lies between clusters whereas 69% is within clusters. These results suggest that between a quarter and a third of the likelihood of experiencing an interruption is related to the cluster that someone lives in, with the remainder due to the specific household, after accounting for the variables within the models.

Table 4. Multi-level logistic regression analysis of user-reported interruptions to piped water services in the urban and peri-urban population of Zambia, 2018 (n = 3047, groups = 177).

	Model 1[a]		Model 2[b]	
Parameter	**Odds Ratio**	**Std.Err.**	**Odds Ratio**	**Std.Err.**
Intercept	0.18***	0.089	0.08***	0.05
Household-level Factors				
Native Language *(ref.: English)*				
Bemba	1.53	0.50	1.88*	0.60
Luvale, Kaonde, Lunda	2.06**	0.74	2.30**	0.81
Lozi	1.81	0.66	1.85	0.66
Other	1.60	0.52	1.77	0.57
Tonga	1.78*	0.62	1.73	0.59
Month of Interview (season) *(ref.: Dec 2018-Jan 2019 (Rainy))*				
Sept 2018 (Dry)	1.64	0.60	2.95***	1.12
Oct-Nov 2018 (Rainy)	1.61	0.53	2.36**	0.80
July-Aug 2018 (Dry)	1.91**	0.61	2.51***	0.84
Type of piped service *(ref.: Piped into dwelling)*				
Piped to yard/plot or Public tap/standpipe	1.20	0.15	1.24	0.16
Piped to neighbour	1.32	0.20	1.33*	0.21
Home ownership *(ref: Does not own)*				
At least partly owns	1.31***	0.13	1.32***	0.13
No information	1.10	0.23	1.10	0.23
Provider-level Dummy Variable				
Provider *(ref.: Lusaka WSC)*				
Chambeshi WSC	2.43**	0.94	-	-
Eastern WSC	0.95	0.40	-	-
Kafubu WSC	8.02***	3.40	-	-
Luapula WSC	2.89**	1.45	-	-
Lukanga WSC	0.36**	0.15	-	-
Mulongo WSC	3.77**	1.85	-	-
Nkana WSC	2.24*	0.90	-	-
North Western WSC	1.39	0.65	-	-
Southern WSC	0.70	0.25	-	-
Western WSC	1.18	0.54	-	-
Household cluster-level Factors				
Neighbourhood Wealth *(ref.: Poor)*				
Middle	-	-	4.31*	3.29
Rich	-	-	3.22	2.31
Provider-level Factors				
Regulator LPCD	-	-	1.01	0.01
GDP per Capita at Service Provider Level	-	-	0.51***	0.11
Regulator LPCD x Neighbourhood Wealth *(ref.: Poor)*				
Middle	-	-	0.97**	0.01
Rich	-	-	0.98*	0.01
Regulator LPCD x GDP per Capita at Service Provider Level	-	-	1.02***	0.01
Random-effects Parameters				
Between PSU variance	1.21	0.19	1.47	0.23
Intraclass Correlation Coefficient (ICC)				
Intra-PSU correlation coefficient	0.27	0.03	0.31	0.03

(Continued)

Table 4. (Continued)

	Model 1[a]		Model 2[b]	
Parameter	**Odds Ratio**	**Std.Err.**	**Odds Ratio**	**Std.Err.**
Log likelihood	-1787.03	-1799.69		
Deviance	3574.00	3599.00		

*** p<0.01

** p<0.05

* p<0.1; PSU: Primary Sampling Unit (household clusters)

[a]Only significant household and household cluster level variables, plus provider dummy variable

[b]Only significant household, household cluster level and provider variables

https://doi.org/10.1371/journal.pwat.0000127.t004

Discussion

Insights into equalities in piped water continuity through integrated analysis of survey, regulator, and provider data sets

Integrated analysis of utility, regulator, and household survey data provides some insights into inequalities and risk factors for piped water continuity in urban Zambia that would not be apparent from analysing any one data set in isolation. There are few statistically significant inequalities at the household level (S17 Table, Model 2) and significant cluster- and provider-level predictors. After controlling for household characteristics, this nationally representative analysis of urban DHS data highlights provider-level inequality in service continuity (Fig 3) [13]. 19% of households reported an interruption in Lukanga, compared to 77% in Kafubu WSC (Table 1), with Kafubu, Mulongo, and several other WSCs having significantly higher odds of household-reported interruptions relative to Lusaka after controlling for household characteristics (Table 4). Similarly, a study [18] comparing DHS with IBNET for Peruvian regions also found high inter-provider and inter-province intermittency, with lower household-reported intermittency in Lima, the capital, than most other regions. Whilst there is some tentative evidence in our models that such provider-level variation in water continuity may relate to utility operational indicators such as LPCD, this variation is also associated with regional socio-economic variation such as in GDP per capita (Table 4).

Alongside provider-level patterns, variation in household-reported water continuity is evident at the household and household cluster level (Tables 2 & 4). Regulatory reports and provider performance databases such as IBNET do not differentiate the level of service provided to different household groups within their PCAs. However, it is known that ethnic minorities, the poor, and other disadvantaged groups may receive poorer quality services [13]. In this analysis, we integrated PCA boundaries with household survey data to examine such inequalities. We find some evidence of urban inequalities in the availability of piped water services in relation to ethnicity (measured via native language), home ownership and seasonality, but not household wealth. We find no significant evidence that such inequalities vary between PCAs. Key findings concerning household- and cluster-level risk factors were as follows:

- We found native English speakers less likely to report interruptions compared to respondents speaking all other native languages. This could reflect the regional distribution of ethnicity in relation to service availability. Ethnicity is an important factor contributing to water inequalities globally, with indigenous populations comprising 15% of the world's poor [1].
- Households in mid-wealth clusters were less likely to experience an interruption than those in poor clusters. This could reflect households in wealthier neighbourhoods adopting

protective measures against interruptions [10, 71] and having greater capacity to address service network failures.

- Counter to our hypothesis, we found households who owned their home were more likely to experience interruptions than those that did not.
- Households using a neighbour's piped water also had higher odds of experiencing an interruption than those with water piped to their dwelling. This could reflect household reluctance to sell their water to neighbours during times of water scarcity.
- As suggested by other authors [18], reported water continuity varied seasonally, with households interviewed in peak rainy season having lower odds of reporting an interruption (Tables 1 and 4).

Although multi-country analyses suggest exposure to microbially contaminated or inadequately chlorinated water is often greater among poorer households [72, 73], we found Zambian households in both poor and wealthy clusters were more exposed to water interruptions relative to those in mid-wealth clusters. Criteria concerning availability and water safety may therefore affect socio-economic inequalities in access to safely managed water services.

Comparability and consistency of survey versus provider and regulatory data

It is clear the data landscape is complex, which is exacerbated by the contrasting approaches to measuring availability [14, 15]. The extent of our analysis has been limited due to the use of two very different metrics of availability: regulator and provider reported yearly mean service hours or LPCD, whilst DHS household survey respondents reported service interruptions in the past fortnight.

Despite this, we find household survey reporting is correlated with provider/regulator reporting (S17 Table). Differences existed between provider and regulator reporting depending on the metric in question, however generally they are highly consistent with one another (S16 Table). This is as expected given NWASCO regulator reports use provider data, but the use of regulator data is preferable for monitoring, given regulator independence from service provision and incentives for providers to report higher service continuity [11].

In linking household survey, regulator and provider data sets to evaluate their consistency, we adopted a different approach to a previous study in Peru [18]. This study aggregated data for PCAs to region level and compared household- and regulator-reported metrics at the region level. In contrast, we used multi-level modelling to assess whether contextual provider- or regulator-reported water continuity metrics were associated with household-level reports of service interruptions. Our approach thus enables simultaneous evaluation of provider-level versus household-level risk factors for supply interruptions, whilst avoiding the known problems of areal aggregation [74]. Particularly for larger service providers, our approach also provides insights into intra-provider socio-economic inequalities in access to an uninterrupted water service via household reports collected independent of the service provider (Table 2).

Limitations affecting Zambia study

This analysis only captures household respondents who report an interruption of at least one full day in the two weeks prior to being surveyed. As such, recall bias will likely exist in the user data [75] whereby respondents may fail to remember or misremember interruptions in their service. For example, respondents using a 2-week recall period systematically under-reported child diarrhoea relative to a 1-week period [76]. Additionally, respondents who spend more time away from their homes may be unaware of short duration outages. Thus, there are likely to be households who experience shorter- and longer-term service interruptions, that are not captured in household survey data. Source of drinking-water and water used for other purposes form inputs to the DHS wealth index, but not reported supply interruptions. This complicates WASH inequality analyses [77] and could have increased collinearity between the wealth index and source type in our models, but we did not find empirical evidence for this.

The quality of provider and regulator data depends greatly on the accuracy of the data reported by each WSC [14]. Data supplied by providers may be of limited reliability as they may lack any form of independent verification [78].

Between datasets, definitional differences existed in the classification of urban/rural areas [31, 34]. Households were excluded based on the DHS urban/rural classification, despite the regulator or provider potentially classifying them as urban. Additionally, given the DHS used a 2010 rural/urban classification [34, 79], it is likely that some rural areas may have become urbanised. The issue of differing classifications is exacerbated by the use of GHSL population data to create cluster-level locality and change in urbanisation variables. GHSL data uses a finer spatial scale classification [80] that differs to the DHS classification. Thus, despite excluding all rural households from the analysis, inconsistencies between datasets are evident. For both variables, some household clusters are classified as rural or unpopulated areas, despite the DHS classifying them as urban. This could also be a result of the DHS's displacement of household cluster locations [35], though the effect is likely minimal as GPS clusters were not displaced across province boundaries and WSCs are coincident with provinces.

The higher-level explanatory variables for wealth also have limitations. The neighbourhood wealth variable was calculated by aggregating household-level DHS wealth index values to household cluster level. This will however mean that closely located extreme differences in wealth are unaccounted for. The provincial wealth variable used provincial GDP data that is based on where industries produce their goods or where their head office is located. Therefore, the data may not be a true representation of wealth in each PCA [65].

Additional explanatory variables were considered for analysis, including 'voice' [9], gender of the person collecting water [81], storage tank ownership [82] and blue water scarcity [83]. In all instances, inclusion was not possible as there was either no available data, limited variability across Zambia or reasonable proxies did not exist.

Barriers to methodological transferability

Our integrated analysis of household survey and water utility databases highlights data-related barriers at present to methodological transferability at both international and national level. In the context of Zambia, comparison has been made possible by a relatively simple data landscape, where PCAs largely coincide with DHS regions. At present, exploring inequalities elsewhere via this methodology would only be possible in countries where household surveys have measured service interruptions and for WSCs serving large populations, with known geographic coverage areas. Given that household survey cluster coordinates are randomly displaced to preserve confidentiality [22], misallocation of households to providers is more likely where PCAs are small. Thus, our approach is more suited to countries with several large-scale water service providers, as opposed to countries dominated by small-scale providers or a single national provider (e.g. Ghana Water Company Limited). Elsewhere, the availability of data from all three streams is currently restricted to eastern and southern Africa (Fig 1 and S1 Fig; S1–S3 Tables), but analysis will become possible in more countries as data availability increases. By 2030, the data landscape may change considerably with greater availability of household survey data on water service continuity, from government via regulators, from utilities as service providers and potentially other new forms of data such as via social media or sensors [84]. Therefore, whilst the current transferability of methods is reliant on present-day data availability, as monitoring expands, the applicability and transferability of this analysis may broaden.

Within Zambia, data were only available for urban and peri-urban services from the provider and regulator, which limited the scope of this analysis to include only the urban and peri-urban third of the Zambian population. Compared to user data which has national coverage, both datasets only include the population served by large-scale reporting commercial utilities and may exclude services from small-scale private schemes/companies [85], micro-operators managing delegated services [86], and community-managed piped services [18].

Implications for international monitoring

The JMP only makes estimates for safely managed drinking-water when there are data available on water quality and at least one other element (accessibility or availability) that represents at least half of the population in question (e.g. country) [87]. The development of methods which could help to reach this criterion threshold, for example by using data from multiple perspectives, will be critical in better analysing WASH for international agendas such as the SDGs. Improvements in data, coupled with a standardised process by which data are processed to give nationally representative and internationally comparable insights into drinking-water availability are needed, especially in order to understand inequalities between population groups [16].

Definitions of provider-reported average service hours requires further clarity. For example, does it represent the hours a pump in a piped network is operated or the average duration of service households receive? Several suggestions have been made to address this, such as regular random household surveys by providers or the use of sensors to detect outages [18]. At present, provider reporting of hrs/day is unclear and it is unknown what the number of days between service is.

We recommend further smaller scale studies of drinking-water availability. Whilst our analysis bears similarities to Rawas *et al.* [18] in its regional/provincial analysis, future work that resonates more closely with Bellaubi *et al.*'s [11] finer scale case study analysis would give more detailed understanding of the availability of piped services. To achieve this, we recommend providers and regulators report availability for smaller geographical units so that households can be better matched with provider jurisdictions. This would also be enhanced by utilities providing their service areas as coverage area boundaries to IBNET to facilitate data integration. Additionally, a future multi-country study could potentially explore other IBNET-derived management indicators (e.g. concerning staff training or adequacy infra-structure investment [88]) alongside those relating to service availability. However, care would be needed to control for other regional covariates that could explain water service continuity, such as regional GDP.

Greater uptake of the JMP's expanded WASH questions by national statistical agencies via household surveys, such as the DHS, would facilitate further, more widespread evaluation of regulator and provider data concerning water availability. Alongside the DHS question analysed in our study, these expanded questions also include asking households *'how many hours per day is water supplied on average*?*'* [2]. Whilst uptake such questions depends on survey implementation resources and national priorities [15], it would enable direct comparison with regulator and provider reporting of service availability via a metric common to all data sources.

Conclusion

This study demonstrates the additional insights into risk factors associated with piped drinking-water availability by incorporating utility and regulatory data into household survey analysis for urban and peri-urban Zambia. At household and household cluster (neighbourhood) level, when adjusting for confounders through multi-level modelling, inequalities are minimal. Wealth was the only neighbourhood risk factor found to influence service availability. At the household level, home ownership, month of interview, native language and type of service had a modest effect. At provider-level, inequalities between PCAs were found in household reporting of interruptions to services.

Our analysis builds on existing assessments of drinking-water services [11, 18] by including the additional perspective of the regulator. We find correlations between user, provider and regulator reports of service availability, but direct comparison is difficult, due to variations in availability metrics used. Limited data availability also restricts more widespread, integrated analyses of these data in rural areas and across SSA. Moving forward, greater availability of water service continuity data from all three data sources should enable assessment of socio-economic and geographic inequalities in access to uninterrupted water services and potential for understanding how water management indicators relate to household-related interruptions in a wider set of countries. However, more widespread, integrated use of utility or regulator data requires investment in government monitoring of intermittency in small-scale community supplies to better understand rural service access. It also requires investment in utility coverage area map layers to facility spatial data integration with household surveys.

Supporting information

S1 Fig. Flowchart showing inclusion criteria for identifying study country, with reasons for excluding African Union member states.
(TIF)

S2 Fig. Scatterplot showing the similarities between 2017 and 2018 regulator reported LPCD.
(TIF)

S3 Fig. Provider inequalities compared to differences in household socioeconomic characteristics, for households that reported experiencing an interruption to their piped water service.
(TIF)

S1 Table. Year of all available data streams for the 13 shortlisted countries. Colour depicts the data stream with dark green = regulator data, mid-green = user data and light green = provider data; black box indicates where data for all three streams is available for the same year.
(TIF)

S2 Table. Metrics of availability used by all available data streams for the 13 shortlisted countries. Colour depicts the data stream with dark green = regulator data, mid-green = user data and light green = provider data; black box indicates where the same metric of availability is used for multiple data streams for a given country.
(TIF)

S3 Table. Level of disaggregation for all available data streams for the 13 shortlisted countries. Colour depicts the data stream with dark green = regulator data, mid-green = user data and light green = provider data; black box indicates where the same metric of availability is

used for multiple data streams for a given country.
(TIF)

S4 Table. Commercial Water Supply Company (WSC) connections and population serviced in 2018.
(TIF)

S5 Table. Outcome and explanatory variables, and their sources, included in the multilevel model analysis.
(TIF)

S6 Table. DHS household and men's survey questions relating to source and continuity of drinking-water supply and socio-economic characteristics.
(TIF)

S7 Table. GDP and GDP per capita based on 2015 current prices, in Zambian Kwacha (ZKW) and US Dollars (US$) for WSC coverage areas.
(TIF)

S8 Table. Proportion of households reporting a full day's interruption in their piped water supply in the preceding fortnight, per provider, with 95% confidence intervals.
(TIF)

S9 Table. User-reported inequalities, by month of interview, in piped-water interruptions by provider.
(TIF)

S10 Table. User-reported inequalities, by type of supply, in piped-water interruptions by provider.
(TIF)

S11 Table. User-reported inequalities, by household size, in piped-water interruptions by provider.
(TIF)

S12 Table. User-reported inequalities, by household wealth, in piped-water interruptions by provider.
(TIF)

S13 Table. User-reported inequalities, by tenure, in piped-water interruptions by provider.
(TIF)

S14 Table. User-reported inequalities, by native language, in piped-water interruptions by provider.
(TIF)

S15 Table. User-reported inequalities in piped-water interruptions by provider: The percentage (weighted) of households, by household cluster characteristics, reporting an interruption to their supply in each PCA.
(TIF)

S16 Table. Assessment of average annual piped water service delivery in urban and peri-urban Zambia in 2017/18 (n = 3047 HHs).
(TIF)

S17 Table. Exploration of clustering in reported service interruptions through comparison of a single-level unconditional model with a two-level variance components model (n = 3047).
(TIF)

S1 Data.
(7Z)

Author Contributions

Conceptualization: Mair L. H. Thomas-Possee, Andrew A. Channon, Robert E. S. Bain, James A. Wright.

Data curation: Mair L. H. Thomas-Possee.

Formal analysis: Mair L. H. Thomas-Possee.

Funding acquisition: Andrew A. Channon, James A. Wright.

Investigation: Mair L. H. Thomas-Possee, Andrew A. Channon, James A. Wright.

Methodology: Andrew A. Channon, James A. Wright.

Project administration: Mair L. H. Thomas-Possee.

Supervision: Andrew A. Channon, Robert E. S. Bain, James A. Wright.

Visualization: Mair L. H. Thomas-Possee.

Writing – original draft: Mair L. H. Thomas-Possee.

Writing – review & editing: Andrew A. Channon, Robert E. S. Bain, James A. Wright.

References

1. UN Water. Sustainable Development Goal 6: Synthesis Report on Water and Sanitation. United Nations. New York, USA; 2018.
2. WHO, UNICEF. Core questions on water, sanitation and hygiene for household surveys: 2018 update. New York, NY, USA.; 2018.
3. Howard G, Bartram J. Domestic Water Quantity, Service Level and Health. World Health Organization. Geneva, Switzerland; 2003.
4. Howard G, Bartram J, Williams A, Overbo A, Fuente D, Geere J-A. Domestic water quantity, service level and health, second edition. World Health Organization. Geneva, Switzerland; 2020.
5. WHO, UNICEF. Progress on household drinking water, sanitation and hygiene 2000–2020: five years into the SDGs. Geneva, Switzerland; 2021.
6. Bivins AW, Sumner T, Kumpel E, Howard G, Cumming O, Ross I, et al. Estimating infection risks and the global burden of diarrheal disease attributable to intermittent water supply using QMRA. Environmental Science and Technology. 2017; 51: 7542–7551. https://doi.org/10.1021/acs.est.7b01014 PMID: 28582618
7. Kumpel E, Nelson KL. Comparing microbial water quality in an intermittent and continuous piped water supply. Water Research. 2013; 47: 5176–5188. https://doi.org/10.1016/j.watres.2013.05.058 PMID: 23866140
8. Lechtenfeld T. Why does piped water not reduce diarrhea for children? Evidence from urban Yemen. Courant Research Centre: Poverty, Equity and Growth-Discussion Papers. Göttingen: Georg-August-Universität Göttingen, Courant Research Centre—Poverty, Equity and Growth (CRC-PEG); 2012.
9. Majuru B, Suhrcke M, Hunter PR. How Do Households Respond to Unreliable Water Supplies? A Systematic Review. International Journal of Environmental Health Research. 2016; 13. https://doi.org/10.3390/ijerph13121222 PMID: 27941695

10. Pattanayak SK, Yang J-C, Whittington D, Bal Kumar KC. Coping with unreliable public water supplies: Averting expenditures by households in Kathmandu, Nepal. Water Resources Research. 2005; 41: 1–11. https://doi.org/10.1029/2003WR002443
11. Bellaubi F, Visscher JT. Water service delivery in Kenya and Ghana: an area-based assessment of water utility performance. Water International. 2014; 39: 952–968. https://doi.org/10.1080/02508060.2015.985976
12. Yu W, Wardrop NA, Bain RES, Lin Y, Zhang C, Wright JA. A Global Perspective on Drinking-Water and Sanitation Classification: An Evaluation of Census Content. Hills RK, editor. PLOS ONE. 2016; 11: e0151645. https://doi.org/10.1371/journal.pone.0151645 PMID: 26986472
13. WHO, UNICEF. Progress on household drinking water, sanitation and hygiene 2000–2017. Special focus on inequalities. New York, USA.; 2019.
14. Majuru B, Suhrcke M, Hunter PR. Reliability of water supplies in low and middle income countries: A structured review of definitions and assessment criteria. Journal of Water, Sanitation and Hygiene for Development. 2018; 8: 142–164. https://doi.org/10.2166/washdev.2018.174
15. Thomas MLH, Channon AA, Bain RES, Nyamai M, Wright JA. Household-Reported Availability of Drinking Water in Africa: A Systematic Review. Water. 2020; 12: 2603. https://doi.org/10.3390/w12092603
16. Bartram J, Brocklehurst C, Fisher M, Luyendijk R, Hossain R, Wardlaw T, et al. Global Monitoring of Water Supply and Sanitation: History, Methods and Future Challenges. International Journal of Environmental Research and Public Health. 2014; 11: 8137–8165. https://doi.org/10.3390/ijerph110808137 PMID: 25116635
17. WHO, UNICEF. Progress on Drinking Water, Sanitation and Hygiene Update and SDG Baselines. Geneva, Switzerland; 2017.
18. Rawas F, Bain R, Kumpel E. Comparing utility-reported hours of piped water supply to households' experiences. npj Clean Water. 2020; 3: 1–9. https://doi.org/10.1038/s41545-020-0053-y
19. DuChanois RM, Liddle ES, Fenner RA, Jeuland M, Evans B, Cumming O, et al. Factors Associated with Water Service Continuity for the Rural Populations of Bangladesh, Pakistan, Ethiopia, and Mozambique. Environmental Science and Technology. 2019; 53: 4355–4363. https://doi.org/10.1021/acs.est.8b07173 PMID: 30917279
20. Smiley SL. Water Availability and Reliability in Dar es Salaam, Tanzania. The Journal of Development Studies. 2016; 52: 1320–1334. https://doi.org/10.1080/00220388.2016.1146699
21. Majuru B, Jagals P, Hunter PR. Assessing rural small community water supply in Limpopo, South Africa: Water service benchmarks and reliability. Science of The Total Environment. 2012; 435: 479–486. https://doi.org/10.1016/j.scitotenv.2012.07.024 PMID: 22885354
22. Perez-Haydrich C, Warren JL, Burgert CR, Emch ME. Guidelines on the use of DHS GPS data. Spatial Analysis Reports No. 8. Calverton, Maryland, USA; 2013 Sep.
23. Alda-Vidal C, Kooy M, Rusca M. Mapping operation and maintenance: an everyday urbanism analysis of inequalities within piped water supply in Lilongwe, Malawi. Urban Geography. 2017; 39: 104–121. https://doi.org/10.1080/02723638.2017.1292664
24. JMP. Downloads Index—Country Files. 2021 [cited 23 Jan 2020]. Available: https://washdata.org/data/downloads#WLD
25. IBNET. The International Benchmarking Network. 2021 [cited 23 Feb 2022]. Available: https://www.ib-net.org/
26. World Bank. Population, total—Zambia. In: Washington D.C., USA [Internet]. 2019 [cited 22 Sep 2020]. Available: https://data.worldbank.org/indicator/SP.POP.TOTL?locations=ZM&view=chart
27. World Bank. Urban population (% of total population)—Zambia. In: Washington D.C., USA [Internet]. 2019 [cited 22 Sep 2020]. Available: https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS?locations=ZM
28. Hamududu BH, Ngoma H. Impacts of climate change on water resources availability in Zambia: implications for irrigation development. Environment, Development and Sustainability. 2020; 22: 2817–2838. https://doi.org/10.1007/s10668-019-00320-9
29. Libanda B, Zheng M, Ngonga C. Spatial and temporal patterns of drought in Zambia. Journal of Arid Land. 2019; 11: 180–191. https://doi.org/10.1007/s40333-019-0053-2
30. Lapworth DJ, Stuart ME, Pedley S, Nkhuwa DCW, Tijani MN. A review of urban groundwater use and water quality challenges in Sub-Saharan Africa. British Geological Survey Open Report OR/17/056; 2017.
31. NWASCO. Urban and Peri-urban Water Supply and Sanitation Sector Report 2018. Lusaka, Zambia: National Water Supply and Sanitation Council; 2018.

32. WHO, UNICEF, JMP. JMP: Zambia Drinking Water Statistics. In: New York, USA [Internet]. 2018 [cited 23 Sep 2020]. Available: https://washdata.org/data/household#!/zmb

33. NWASCO. Annual Report 2018. Lusaka, Zambia: National Water Supply and Sanitation Council; 2018.

34. Zambia Statistics Agency M of H (MOH) Z, ICF. Zambia Demographic and Health Survey 2018. Lusaka, Zambia, and Rockville, Maryland, USA; 2019.

35. Burgert CR, Colston J, Roy T, Zachary B. Geographic displacement procedure and georeferenced data release policy for the Demographic and Health Surveys. DHS Spatial Analysis Reports No. 7. ICF International, Calverton, Maryland, USA; 2013.

36. IBNET. Country Profile Zambia. 2020 [cited 23 Sep 2020]. Available: https://database.ib-net.org/country_profile?ctry=29&years=2019,2018,2017,2016,2015&type=report&ent=country&mult=true&table=true&chart=false&chartType=column&lang=en&exch=1

37. IBNET. Zambia Utility Company Continuity of Supplies. 2017 [cited 12 Oct 2020]. Available: https://database.ib-net.org/utilities_results?uid=22094,20,22095,125,24382,22099,121,122,123,22096,126,22097,127,22098,124&years=2017&type=indicator&ent=utility&mult=true&report=1&indicator=34&table=true&chart=true&chartType=column&lang=EN&exch=1

38. IBNET. Zambia Utility Company Total Water Consumption (LPCD). 2017 [cited 12 Oct 2020]. Available: https://database.ib-net.org/utilities_results?uid=22094,20,22095,125,24382,22099,121,122,123,22096,126,22097,127,22098,124&years=2017&type=indicator&ent=utility&mult=true&report=1&indicator=7&table=true&chart=true&chartType=column&lang=EN&exch=1

39. IBNET. Non revenue water. 2021 [cited 21 Oct 2020]. Available: https://www.ib-net.org/toolkit/ibnet-indicators/non-revenue-water/

40. Simukonda K, Farmani R, Butler D. Intermittent water supply systems: causal factors, problems and solution options. Urban Water Journal. 2018; 15: 488–500. https://doi.org/10.1080/1573062X.2018.1483522

41. Liemberger R, Wyatt A. Quantifying the global non-revenue water problem. Water Science and Technology: Water Supply. 2019; 19: 831–837. https://doi.org/10.2166/ws.2018.129

42. NWASCO. Urban and Peri-Urban WSS Sector Reports. In: National Water Supply and Sanitation Council [Internet]. 2020 [cited 23 Sep 2020]. Available: http://www.nwasco.org.zm/index.php/media-center/publications/urban-and-peri-urban-wss-sector-reports

43. National Water Supply and Sanitation Council. Water Sector Overview. In: National Water Supply and Sanitation Council [Internet]. 2023 [cited 22 Mar 2023]. Available: https://www.nwasco.org.zm/index.php/regulated-sector/water-sector-overview

44. GADM. Zambia Administrative Boundaries. 2020 [cited 23 Sep 2020]. Available: https://gadm.org/download_country_v3.html

45. Croft TN, Marshall AMJ, Allen CK. Guide to DHS Statistics: DHS-7. Maryland, USA; 2018.

46. Luh J, Bartram J. Drinking water and sanitation: progress in 73 countries in relation to socioeconomic indicators. Bulletin of the World Health Organization. 2016; 94: 111–121. https://doi.org/10.2471/BLT.15.162974 PMID: 26957676

47. Cassivi A, Guilherme S, Bain R, Tilley E, Owen E, Waygood D, et al. Drinking water accessibility and quantity in low and middle-income countries: A systematic review. International Journal of Hygiene and Environmental Health. 2019; 222: 1011–1020. https://doi.org/10.1016/j.ijheh.2019.06.011 PMID: 31320308

48. Rutstein SO. The DHS Wealth Index: Approaches for Rural and Urban Areas. Calverton, Maryland, USA; 2008.

49. Zoungrana TD. The effect of wealth on the choice of household drinking water sources in West Africa. International Journal of Finance & Economics. 2020; ijfe.1903. https://doi.org/10.1002/ijfe.1903

50. Kayaga S, Franceys R. Costs of urban utility water connections: Excessive burden to the poor. Utilities Policy. 2007; 15: 270–277. https://doi.org/10.1016/j.jup.2007.06.002

51. Dungumaro EW. Socioeconomic differentials and availability of domestic water in South Africa. Physics and Chemistry of the Earth. 2007; 32: 1141–1147.

52. Kelly E, Shields KF, Cronk R, Lee K, Behnke N, Klug T, et al. Seasonality, water use and community management of water systems in rural settings: Qualitative evidence from Ghana, Kenya, and Zambia. Science of the Total Environment. 2018; 628–629: 715–721. https://doi.org/10.1016/j.scitotenv.2018.02.045 PMID: 29454211

53. Simukonda K, Farmani R, Butler D. Causes of intermittent water supply in Lusaka City, Zambia. Water Practice and Technology. 2018; 13: 335–345. https://doi.org/10.2166/wpt.2018.046

54. International Research Institute of Climate and Society, World Bank. World Bank Climate Variability Tool: Precipitation Plots. 2021 [cited 12 May 2021]. Available: http://iridl.ldeo.columbia.edu/maproom/Global/World_Bank/Climate_Variability/

55. Jackson K. Diversity and the Distribution of Public Goods in Sub-Saharan Africa. Journal of African Economics. 2013; 22: 437–462.

56. Arouna A, Dabbert S. Determinants of domestic water use by rural households without access to private improved water sources in Benin: A Seemingly Unrelated Tobit Approach. Water Resources Management. 2010; 24: 1381–1398. https://doi.org/10.1007/s11269-009-9504-4

57. Armand L, Fotue T, Sikod F. Determinants of the households' choice of drinking water source in Cameroon. Journal of Sustainable Development in Africa. 2012; 14: 86–97.

58. Mulenga JN, Bwalya BB, Kaliba-Chishimba K. Determinants and inequalities in access to improved water sources and sanitation among the Zambian households. International Journal of Development and Sustainability. 2017; 6: 746–762.

59. Staddon C, Rogers J, Warriner C, Ward S, Powell W. Why doesn't every family practice rainwater harvesting? Factors that affect the decision to adopt rainwater harvesting as a household water security strategy in central Uganda. Water International. 2018; 43: 1114–1135. https://doi.org/10.1080/02508060.2018.1535417

60. Pesaresi M, Freire S. GHS-SMOD R2016A - GHS settlement grid, following the REGIO model 2014 in application to GHSL Landsat and CIESIN GPW v4-multitemporal (1975-1990-2000-2015). In: European Commission, Joint Research Centre (JRC) [Internet]. 2016 [cited 12 Oct 2020]. Available: https://data.jrc.ec.europa.eu/dataset/jrc-ghsl-ghs_smod_pop_globe_r2016a

61. Winter S, Dreibelbis R, Barchi F. Context matters: a multicountry analysis of individual- and neighbourhood-level factors associated with women's sanitation use in sub-Saharan Africa. Tropical Medicine & International Health. 2018; 23: 173–192. https://doi.org/10.1111/tmi.13016 PMID: 29172229

62. Robles-Velasco A, Cortés P, Muñuzuri J, Onieva L. Prediction of pipe failures in water supply networks using logistic regression and support vector classification. Reliability Engineering and System Safety. 2020; 196: 106754. https://doi.org/10.1016/j.ress.2019.106754

63. Kennedy-Walker R, Amezaga JM, Paterson CA. The impact of community social dynamics on achieving improved sanitation access for the urban poor: The case of Lusaka, Zambia. Habitat Int. 2015; 50: 326–334. https://doi.org/10.1016/J.HABITATINT.2015.09.004

64. Felice E. The Misty Grail: The Search for a Comprehensive Measure of Development and the Reasons for GDP Primacy. Development and Change. 2016; 47: 967–994. https://doi.org/10.1111/dech.12257

65. Republic of Zambia Central Statistical Office. Research Paper on Provincial Gross Domestic Product. Lusaka, Zambia; 2017.

66. Republic of Zambia Central Statistical Office. Zambia in Figures 2018. Lusaka, Zambia; 2018.

67. StataCorp. Stata Statistical Software: Release 16. College Station, TX: StataCorp LLC.; 2019.

68. Pfeffermann D, Skinner CJ, Holmes DJ, Goldstein H, Rasbash J. Weighting for unequal selection probabilities in multilevel models. Journal of the Royal Statistical Society Series B: Statistical Methodology. 1998; 60: 23–40. https://doi.org/10.1111/1467-9868.00106

69. Snijders TAB. Power and sample size in multilevel modeling. In: Everitt BS, Howell DC, editors. Encyclopedia of Statistics in Behavioral Science. Chichester, UK: Wiley; 2005. pp. 1570–1573.

70. Elkasabi M, Ren R, Pullum TW. DHS Methodology Reports 27. Multilevel Modeling Using DHS Surveys: A Framework to Approximate Level-Weights. Rockville, Maryland, USA; 2020.

71. Oageng I, Power Mmopelwa G. Water consumption patterns in a rural setting in Ngamiland district, Botswana: the case of the Boro village. Journal of Water Sanitation and Hygiene for Development. 2014; 4: 720–726. https://doi.org/10.2166/washdev.2014.065

72. Yang H, Bain R, Bartram J, Gundry S, Pedley S, Wright J. Water Safety and Inequality in Access to Drinking-water between Rich and Poor Households. Environmental Science & Technology. 2013; 47: 11222–1230. https://doi.org/10.1021/es303345p PMID: 23276231

73. Bain R, Johnston R, Khan S, Hancioglu A, Slaymaker T. Monitoring Drinking Water Quality in Nationally Representative Household Surveys in Low- and Middle-Income Countries: Cross-Sectional Analysis of 27 Multiple Indicator Cluster Surveys 2014–2020. Environmental Health Perspectives. 2021; 129: 097010. https://doi.org/10.1289/EHP8459 PMID: 34546076

74. Nelson JK, Brewer CA. Evaluating data stability in aggregation structures across spatial scales: revisiting the modifiable areal unit problem. Cartography and Geographic Information Science. 2017; 44: 35–50.

75. Boerma JT, Sommerfeltb AE. Demographic and health surveys (DHS): contributions and limitations. World health statistics quarterly. 1993; 46: 222–226. PMID: 8017081

76. Overbey KN, Schwab KJ, Exum NG. Comparison of 1-week and 2-week recall periods for caregiver-reported diarrhoeal illness in children, using nationally representative household surveys. International Journal of Epidemiology. 2019; 48: 1228–1239. https://doi.org/10.1093/ije/dyz043 PMID: 30907423

77. Rheingans R, Anderson JD, Luyendijk R, Cumming O. Measuring disparities in sanitation access: does the measure matter? Tropical Medicine & International Health. 2014; 19: 2–13. https://doi.org/10.1111/tmi.12220 PMID: 24851256

78. UNICEF. Drinking Water: Equity, safety and sustainability. In: United Nations International Children's Emergency Fund. New York, USA [Internet]. UNICEF/WHO. Geneva, Switzerland; 2021 [cited 14 Jun 2021]. Available: https://data.unicef.org/resources/drinking-water-equity-safety-and-sustainability-2011-thematic-report/

79. Elkasabi M, Thomas MLH. Zambia DHS 2018: Classification of rural/urban. In: The DHS Program User Forum [Internet]. 2020 [cited 15 Mar 2023]. Available: https://userforum.dhsprogram.com/index.php?t=msg&goto=20391&&srch=zambia+rural#msg_20391

80. Dijkstra L, Poelman H. A harmonised definition of cities and rural areas: the new degree of urbanisation Working Papers. Brussels, Belgium; Luxembourg City, Luxembourg; 2014.

81. Jeil EB, Abass K, Ganle JK. "We are free when water is available": Gendered livelihood implications of sporadic water supply in Northern Ghana. Local Environment. 2020; 25: 320–335. https://doi.org/10.1080/13549839.2020.1744118

82. Guragai B, Takizawa S, Hashimoto T, Oguma K. Effects of inequality of supply hours on consumers' coping strategies and perceptions of intermittent water supply in Kathmandu Valley, Nepal. Science of The Total Environment. 2017; 599–600: 431–441. https://doi.org/10.1016/j.scitotenv.2017.04.182 PMID: 28482301

83. McNally A, Verdin K, Harrison L, Getirana A, Jacob J, Shukla S, et al. Acute Water-Scarcity Monitoring for Africa. Water. 2019; 11: 1968. https://doi.org/10.3390/W11101968

84. Thomson P, Hope R, Foster T. GSM-enabled remote monitoring of rural handpumps: A proof-of-concept study. Journal of Hydroinformatics. 2012; 14: 829–839. https://doi.org/10.2166/hydro.2012.183

85. NWASCO. Water Providers. In: National Water Supply and Sanitation Council [Internet]. Lusaka, Zambia; 2021 [cited 4 Aug 2021]. Available: http://www.nwasco.org.zm/index.php/regulated-sector/water-providers

86. Castro V, Morel A. Can delegated management help water utilities improve services to informal settlements? Waterlines. 2008; 27: 289–306.

87. WHO, UNICEF. WASH in the 2030 Agenda New global indicators for drinking water, sanitation and hygiene. Geneva, Switzerland; 2017.

88. IBNET. IBNET Indicators. 2023 [cited 14 Mar 2023]. Available: https://www.ib-net.org/toolkit/ibnet-indicators/