

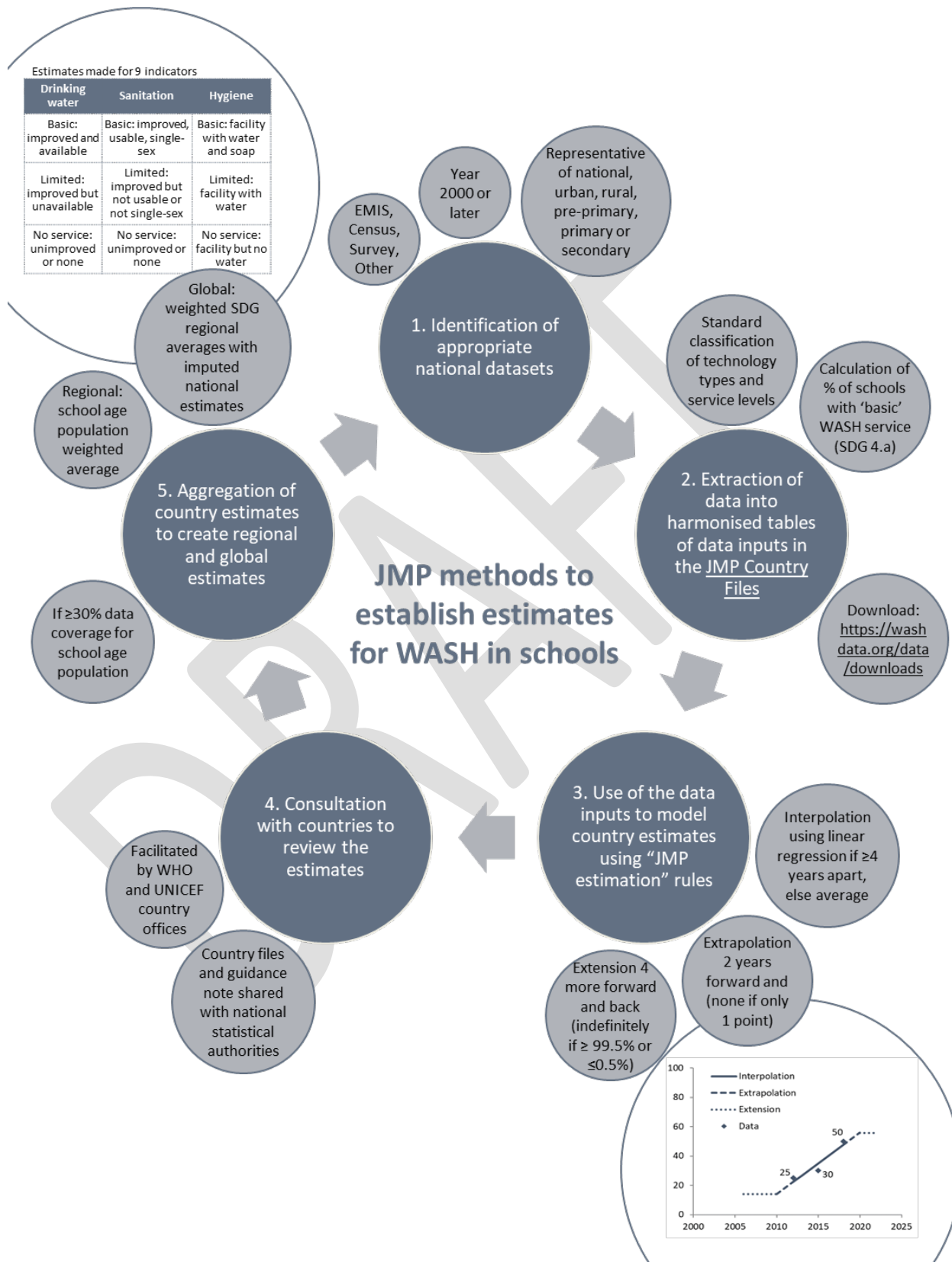
JMP Methodology for WASH in Schools

November 2021

Contents

Graphical summary	1
Introduction	2
Background	2
0. Method overview	3
1. Identification of national datasets	4
2. Data extraction, cleaning and pre-processing	5
2.1 Data compilation	6
2.2 Technology classifications	7
2.2.1 Drinking water	7
2.2.2 Sanitation	7
2.2.3 Hygiene	8
2.3 Service level classifications	8
2.3.1 Drinking water	8
2.3.2 Sanitation	9
2.3.3 Hygiene	11
2.4 Data cleaning and adjustment	12
2.5 Data acceptance	13
3. Producing country estimates	14
3.1 National, rural and urban, and school level estimates	14
3.2 Primary indicators	14
3.2.1 Adjustments	17
3.3 Annual Rates of Change	17
4 Country consultation	18
5 Producing regional and global estimates	19
5.1 School age population	19
5.2 Service level estimates	20
Annex 1: JMP Country Files for WASH in schools	22
Annex 2: Data flows in the JMP country file for WASH in Schools	24

Graphical summary



Introduction

The WHO/UNICEF Joint Monitoring Programme for Water Supply, Sanitation and Hygiene (JMP) was established in 1990 and has developed and refined methods for global monitoring of drinking water, sanitation and hygiene (WASH) in households¹. In 2016, the JMP developed methods for global monitoring of WASH in schools to support the establishment of baseline estimates for the relevant Sustainable Development Goal (SDG) targets (4.a, 6.1, and 6.2) and to track progress over time.²

The purpose of this document is to explain the methods used by the JMP to compile and harmonize available data on “basic” drinking water, sanitation and hygiene services in schools and establish comparable national estimates for SDG monitoring. This document describes the methods and rules applied in the JMP baseline report and first progress update on WASH in schools published in August 2018³ and September 2020⁴, respectively. As with global monitoring of WASH in other settings, JMP methods for estimating progress on WASH in schools will continue to be refined over time.

Background

The terms “universal” and “for all” in Targets 6.1 and 6.2 highlight the need for monitoring WASH beyond the household, including institutions such as schools (**Table 1**). Target 4.a includes WASH in schools, explicitly, with an associated indicator of the “proportion of schools with access to:...(e) basic drinking water; (f) single-sex basic sanitation; and (g) basic handwashing facilities” (**Table 1**). WHO and UNICEF through the JMP are the custodian agencies for indicators related to targets 6.1 and 6.2, while UNESCO-UIS is the custodian agency for indicator 4.a.1. JMP estimates on WASH in schools are used within the WASH sector as well as shared with UNESCO-UIS in support of SDG 4.a monitoring.

Table 1. SDG targets and indicators related to WASH in schools

Goals	Targets	Indicators
6: Ensure availability and sustainable management of water and sanitation for all	6.1: By 2030, achieve universal and equitable access to safe and affordable drinking water for all	See 4.a.1 for WASH in schools indicators
	6.2: By 2030 achieve access to adequate and equitable sanitation and hygiene for all and end open defecation, paying special attention to the needs of women and girls and those in vulnerable situations	See 4.a.1 for WASH in schools indicators
4: Ensure inclusive & equitable quality education & promote lifelong learning opportunities for all	4.a: Build and upgrade education facilities that are child, disability and gender sensitive and provide safe, non-violent, inclusive and effective learning environments for all	4.a.1. Proportion of schools with access to: (a) electricity; (b) the Internet for pedagogical purposes; (c) computers for pedagogical purposes; (d) adapted infrastructure and materials for students with disabilities; (e) basic drinking water; (f) single-sex basic sanitation facilities; and (g) basic handwashing facilities (as per the Water, Sanitation and Hygiene for All indicator definitions) ⁵

¹ WHO and UNICEF. 2018. JMP methodology: 2017 update and SDG baselines.

<<https://washdata.org/sites/default/files/documents/reports/2018-04/JMP-2017-update-methodology.pdf>>

² WHO and UNICEF. 2018. Core questions and indicators for monitoring WASH in schools in the Sustainable Development Goals.

<<https://washdata.org/sites/default/files/documents/reports/2018-08/SDGs-monitoring-wash-in-schools-2018-August-web2.pdf>>

³ WHO and UNICEF. 2018. Drinking water, sanitation and hygiene in schools: global baseline report 2018.

<<https://washdata.org/sites/default/files/documents/reports/2018-11/JMP%20WASH%20in%20Schools%20WEB%20final.pdf>>

⁴ WHO and UNICEF. 2020. Progress on drinking water, sanitation and hygiene in schools: special focus on COVID-19.

<<https://washdata.org/sites/default/files/2020-09/JMP-2020-WASH-schools.pdf>>

⁵ This refers to the definitions presented in this document.

The WASH-related sub-indicators include:

- Proportion of schools with access to basic drinking water, by education level (%);
- Proportion of schools with access to access to single-sex basic sanitation, by education level (%); and
- Proportion of schools with basic handwashing facilities, by education level (%)

Similar to JMP monitoring of household WASH, multi-level service ladders for monitoring WASH in schools (**Figure 1**) enable countries at different stages of development to track and compare progress in reducing inequities. There are separate ladders for drinking water, sanitation and hygiene services. Within each category, the service ladder includes three levels: no service, limited service, and basic service, where the *basic* service threshold corresponds to the SDG indicator for Target 4.a.⁶

SERVICE LEVEL	DRINKING WATER	SANITATION	HYGIENE
BASIC SERVICE	Drinking water from an improved source and water is available at the school at the time of the survey	Improved sanitation facilities at the school that are single-sex and usable (available, functional and private) at the time of the survey	Handwashing facilities with water and soap available at the school at the time of the survey
LIMITED SERVICE	Drinking water from an improved source but water is unavailable at the school at the time of the survey	Improved sanitation facilities at the school that are either not single-sex or not usable at the time of the survey	Handwashing facilities with water but no soap available at the school at the time of the survey
NO SERVICE	Drinking water from an unimproved source or no water source at the school	Unimproved sanitation facilities or no sanitation facilities at the school	No handwashing facilities or no water available at the school

Figure 1. JMP service ladders for global monitoring of WASH in schools

0. Method overview

The JMP produces estimates for a total of nine primary indicators related to drinking water, sanitation and hygiene in schools. Estimates are produced at the country, regional, and the global levels. All estimates are expressed as the proportion of schools with a type of drinking water, sanitation, and hygiene technology or level of service (**Table 2**) and are reported as a proportion of all schools, not ‘of which’ (e.g. the proportion with *improved* toilets is out of the total number of schools surveyed, not out of those with a toilet).

Table 2. Nine primary indicators produced by the JMP for WASH in schools

Water	The proportion of schools with...
W ₁	any drinking water source
W ₂	improved ⁷ drinking water sources
W ₃	water from an improved drinking water source that is available at the school at the time of the survey (basic drinking water service - SDG indicator 4.a.1 ⁸)
Sanitation	The proportion of schools with...
S ₁	any sanitation facility
S ₂	improved ⁷ sanitation facilities

⁶ For countries where “basic” service is not an aspirational aim and further monitoring resources are available, an additional “advanced” service level is suggested. These are not currently monitored at the global level, but the JMP highlights examples in progress reports.

⁷ See section 2.2 for an explanation of “improved” facilities

⁸ Metadata of global indicator for SDG target 4.a.1: <https://unstats.un.org/sdgs/metadata/files/Metadatas-04-0A-01.pdf>

S ₃	single-sex improved sanitation facilities that are usable (available, functional, private) at the time of the survey (basic sanitation service - SDG indicator 4.a.1)
Hygiene	The proportion of schools with...
H ₁	any handwashing facility
H ₂	handwashing facilities with water
H ₃	handwashing facilities with water and soap available at the time of the survey (basic hygiene service - SDG indicator 4.a.1)

The service level indicators are then calculated on the basis of the relevant primary indicators, as shown in **Table 3**. In some cases data sources report the basic service level directly, without disaggregation into the sub-indicators (primary indicators). The “limited service” and “no service” level indicators are calculated by simple arithmetic operations on primary indicators.

Table 3. 9 service level indicators (secondary indicators) produced by the JMP for WASH in schools

Water	The proportion of schools with...	Means of calculation
W ₃	water from an improved drinking water source that is available at the school at the time of the survey (basic drinking water service - SDG indicator 4.a.1 ⁹)	Primary regression of W ₃
W ₄	limited water service	$W_2 - W_3$
W ₅	no water service	$100 - W_2$
Sanitation	The proportion of schools with...	
S ₃	single-sex improved sanitation facilities that are usable (available, functional, private) at the time of the survey (basic sanitation service - SDG indicator 4.a.1)	Primary regression of S ₃
S ₄	Limited sanitation service	$S_2 - S_3$
S ₅	No sanitation service	$100 - S_2$
Hygiene	The proportion of schools with...	
H ₃	handwashing facilities with water and soap available at the time of the survey (basic hygiene service - SDG indicator 4.a.1)	Primary regression of H ₃
H ₄	Limited hygiene service	$H_2 - H_3$
H ₅	No hygiene service	$100 - H_2$

Wherever possible, these estimates are disaggregated by urban and rural, and by school level, including pre-primary, primary, and secondary.

Production of the estimates follows a consistent series of steps, including:

1. Identification of appropriate national datasets,
2. Extraction of data from national datasets into harmonised tables of data inputs after data cleaning and pre-processing,
3. Use of the data inputs to model country estimates,
4. Consultation with countries to review the estimates, and
5. Aggregation of country estimates to create regional and global estimates.

This document describes each of these steps in detail.

1. Identification of national datasets

The JMP has developed a global database on WASH in schools to support monitoring of the SDGs which is populated with national datasets. The JMP team conducts regular data searches by systematically visiting the websites of national statistical offices, and key sector institutions such as

⁹ Metadata of global indicator for SDG target 4.a.1: <https://unstats.un.org/sdgs/metadata/files/Metadatas-04-0A-01.pdf>

ministries of water and sanitation, education and health, and regulators of drinking water and sanitation services. Other regional and global databases are also reviewed for new datasets, including Eurostat, the *UNESCO Institute for Statistics* (UIS), and international organizations implementing facility surveys such as the *World Bank Service Delivery Indicators* (SDI) surveys, and the *UNESCO Latin American Laboratory for Assessment of the Quality of Education* (LLECE). UNICEF and WHO regional and country offices provide support to identify newly available school surveys and censuses, including national *Education Management Information Systems* (EMIS).

For global monitoring of the SDGs, the JMP restricts analysis to data from 2000 onwards. Datasets included in the JMP global database on WASH in schools include:

- **National EMIS**, which in principle collect basic information from all schools within a country typically on an annual basis. EMIS are managed by ministries of education and include data from public schools and, in some countries, private, religious and community-run schools.
- **Periodic school censuses**, which aim to collect basic information from all schools in a country but are administered irregularly and outside of the EMIS.
- **School surveys**, which collect detailed information from a subset of schools.¹⁰ These may target national, rural, or urban schools, or a specific school level between pre-primary, primary and secondary. An appropriate sample design is necessary for survey results to be representative, and surveys are often led by or reviewed and approved by national statistical organizations or ministries of education.
- **Other**, which are secondary data sources reported through third parties, including UIS which compile data collected through education questionnaires that national authorities are requested to complete on an annual basis. These and other secondary data sources are only used to calculate JMP country estimates when primary national datasets with supporting documentation are unavailable.

Some datasets reviewed by the JMP are not representative of national, rural or urban schools, or specific school levels. The JMP includes datasets in its database when they include data from at least 50 schools in a relevant domain (e.g. national, urban, rural, pre-primary, primary, or secondary), or if they represent at least 50% of schools in a relevant domain. School censuses with at least one-third (or a 33%) response rate are included. However, not all datasets in the database are used in the production of estimates (see section 2.6, Data Acceptance). For example, the JMP database used to produce the 2020 update report included 1029 datasets, of which 867 were used to generate country estimates. These include 814 datasets used to generate estimates at the national level as well as datasets that were only representative of a particular domain such as rural areas or pre-primary schools.

2. Data extraction, cleaning and pre-processing

Once datasets are identified, national data on drinking water, sanitation and hygiene in schools are extracted from the original sources and recorded in standardized data extraction templates. All available data for each country, area, or territory are compiled in individual spreadsheets (the JMP country files), with drinking water, sanitation, and hygiene data recorded on separate sheets (see

¹⁰ Examples include international survey programmes such as the SDI and LLECE, and national survey programmes such as the *Bangladesh National Hygiene Survey*.

Annex 1 for details). The JMP global databases include information for a total of 234 countries, areas, and territories included in the UN Population Division's World Population Prospects tables however not all of these have data on WASH in schools. Country files for all countries with national data available on WASH in schools can be downloaded from the JMP website.¹¹

2.1 Data compilation

The JMP compiles data from a wide range of data sources and formats. Where possible, the JMP uses original microdata to produce its own tabulations, which are then transferred to the country files. However, in many cases microdata are not readily accessible and, in such cases, relevant data are transcribed from reports available in different formats (e.g. PDFs, Word files, Excel spreadsheets, etc...).

Year assignment

Datasets are assigned a single year in the JMP database, even when data collection and publication span several years. The year assigned to a dataset is the last year in which data collection was made or the ending year of the academic year during which the survey was conducted. Thus, a survey administered from 1 October 2021 through 2 January 2022 would be assigned to 2022, and an EMIS conducted during the 2021-2022 academic year would be assigned to 2022.

Some surveys and administrative data sources involve multiple rounds of data collection within a single year. In such cases, data will be combined into a single set of statistics representing average values for the year, except for specific cases where national authorities recommend using the minimum value due to large seasonal variations in WASH service levels.

Geographic scope

The JMP country files include data that are representative of all schools nationally as well as data that are representative of urban schools and rural schools, or representative of schools at specific levels, including pre-primary, primary or secondary. Where possible, data for all six domains are extracted but, in some cases, data are either unavailable or it is not possible to disaggregate them for urban and rural areas or by school level. Data that are not representative of any of these domains are generally not extracted into the country files. In cases where recent representative data sources are scarce, non-representative data sources may be inserted into the country files but flagged as non-representative and not used to generate estimates.

Weighting

If national data are unavailable, but data are reported for urban and rural schools, or specific school levels, a national estimate is calculated based on a weighted average from other domains. Data are weighted by the total number of schools in each domain, when that information is available (e.g. from the national EMIS). When the number of schools is unknown, UIS data on distribution of school age population are used to calculate a weighted average of data representative of specific school levels, and UN Population Division estimates of the proportion of population living in urban areas¹² are used to calculate a weighted average of data representative of urban and rural schools.

School level classification

¹¹ <https://washdata.org/data/downloads>

¹² UN DESA World Population Prospects (<https://population.un.org/wpp/>) and World Urbanization Projects (<https://population.un.org/wup/>) may differ from national population estimates.

Unless otherwise categorised by national authorities, all schools with pre-primary level students are classified as ‘pre-primary’¹³, all schools with primary level students are classified as ‘primary’, and all schools with secondary level students are classified as ‘secondary’. This means that some schools may be double- or triple-counted and the total number of schools does not necessarily equal the sum of the pre-primary, primary and secondary schools.

2.2 Technology classifications

The JMP uses a standard international classification system to categorise populations using different drinking water, sanitation and hygiene technologies. Data extracted from national datasets are matched to the corresponding standard JMP classifications to enable comparison between data sources within and between countries. Technologies are classified based on original definitions, questionnaires and guidance provided for each data source, and reference to other data sources already on file.

2.2.1 Drinking water

Drinking water sources used by students may be located on or off the school premises and are classified as either “improved” or “unimproved” based on technology (**Table 3**). “Improved” drinking water sources are those that have the potential to deliver safe water by nature of their design and construction.¹⁴ The use of “improved” drinking water sources served as the main Millennium Development Goal (MDG) indicator on drinking water in households and forms the foundation of the new SDG indicators on drinking water in schools. Schools that rely on students to bring drinking water from home are classified as having no water source.

Table 4. Classification of drinking water technologies in schools

Facility classification	Improved	Unimproved
Piped water supply	X	
Tubewell, borehole	X	
Protected well/spring	X	
Unprotected well/spring		X
Rainwater	X	
Packaged water* (bottled or sachet)	X	
Delivered water* (tanker-truck or cart)	X	
Surface water (river, lake, dam, pond, stream, irrigation channel)		X
Other improved	X	
Other unimproved		X
No water source	N/A	N/A

*Packaged and delivered water were previously categorised as unimproved for MDG monitoring

2.2.2 Sanitation

Sanitation facilities used by students may be located inside the school building or within the school compound and are classified as either “improved” or “unimproved” based on technology (**Table 4**). While classification terminology varies for different settings, in general, “improved” sanitation facilities are those designed to hygienically separate excreta from human contact.¹⁴ The use of “improved” sanitation facilities served as the main Millennium Development Goal (MDG) indicator

¹³ Where data are available for Early Childhood Development centres these are classified as ‘pre-primary’.

¹⁴ See washdata.org for more information on the JMP definitions for "improved" facilities, as well as current categorizations.

on sanitation in households and forms the foundation of the new SDG indicators on sanitation in schools.

Table 5. Classification of sanitation technologies in schools

Facility classification	Improved	Unimproved
Flush toilets	X	
Pour-flush toilets	X	
Ventilated Improved Pit (VIP) latrines	X	
Pit latrines with slab*	X	
Pit latrines without slab/open pit		X
Composting toilets	X	
Hanging toilets/hanging latrines		X
Bucket latrines		X
Other improved	X	
Other unimproved		X
No latrines (bush, field, ditch)	N/A	N/A

*A slab is platform constructed from materials that are durable and easy to clean that covers the pit and has a hole that allows excreta to be deposited directly into the pit. Pit latrines that are partially covered or with materials that are not durable and easy to clean (e.g. sticks, logs or bamboo) are classified as “pit latrines without slab”.

2.2.3 Hygiene

“Hygiene” can include many aspects, for example, oral hygiene, menstrual hygiene, or food hygiene. For the purposes of global monitoring of WASH in schools, basic “hygiene” service refers to hand hygiene, and in particular to the availability of facilities for handwashing at school.

Handwashing facilities used by students may be located inside the school building or within the school compound. They may be fixed or mobile and include a sink with tap water, buckets with taps, tippy-taps, and jugs or basins designated for handwashing.

2.3 Service level classifications

Additional data are used to calculate the proportion of schools that meet the service level criteria of the SDG global indicators. The following sub-sections outline the criteria used to classify schools as having ‘basic’, ‘limited’ or ‘no’ service based on available data and the assumptions used when national datasets are incomplete.

It is recognized that the SDG criteria do not consider all elements of the human rights to safe water and sanitation and, where appropriate, some countries may choose to monitor additional elements such as drinking water quality, or accessibility of toilets for students with limited mobility. While additional criteria are not yet monitored at the global level due to limited availability of national data, country level examples are collected and highlighted by the JMP and may be considered for systematic global monitoring in the future.²

2.3.1 Drinking water

In addition to the type of drinking water source (technology classification) used by the school, the service level classification requires information on the **availability** of drinking water at the school.

Availability

Drinking water is considered to be “available” if water is available at the school at the time of the survey or questionnaire, either directly from the main water supply, or water collected from the main water supply and stored at the school. In cases where schools report water availability on

average rather than the day of the survey, schools with water available at least 50% of the time (e.g. at least 4 hours per 8-hour school day or 3 days per 5-day school week) are classified as having drinking water available. In the absence of data on water availability, the functionality of the water system is used as a proxy where water facilities classified as ‘functional’ are assumed to have water available.

Calculation of basic drinking water service

The proportion of schools with basic drinking water services is calculated as the proportion of schools with drinking water available at the school from an improved source (**Figure 2**). Schools with an improved drinking water source but water not available are classified as having limited service. The remaining proportion of schools with water from an unimproved source, or no water source at all, are classified as having no drinking water service. In cases where there are data on any source but no data on improved sources from any dataset, no service is estimated based on the proportion of schools with no water source.

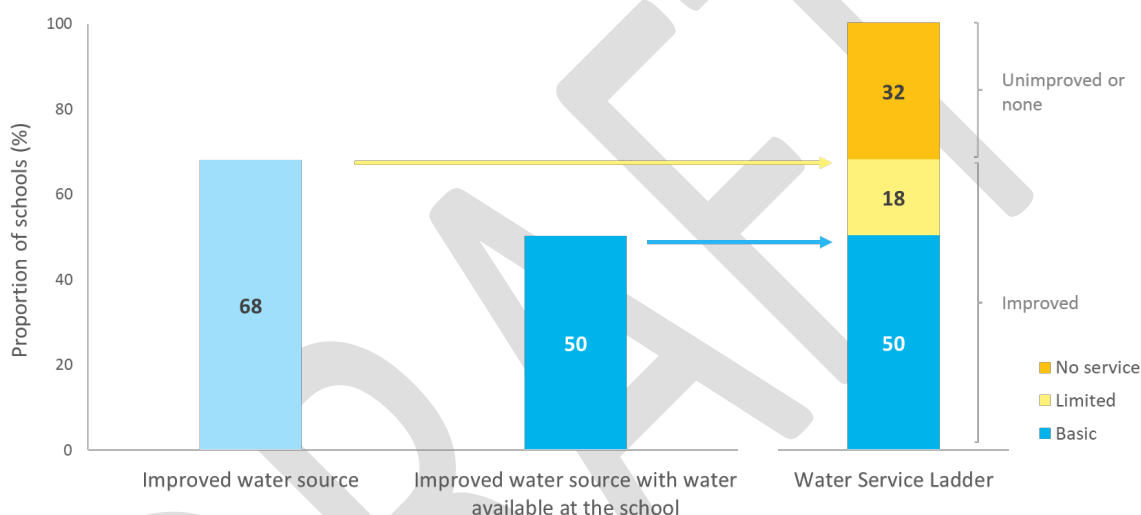


Figure 2. Example of calculating basic drinking water service in schools

In cases where microdata are unavailable, but the proportion of schools with an improved water source is known and the proportion of schools with water available is known, basic service is calculated by multiplying the proportion of schools with an improved water source by the ratio of schools with water available to schools with some type of water supply. For example, if 60% of schools have water available, 80% have water from an improved supply, and 5% have no water supply, it is estimated that 50.5% of schools have basic drinking water service ($60/95 \times 80 = 50.5$). Estimates based on incomplete datasets are reviewed as new data become available.

2.3.2 Sanitation

In addition to the type of sanitation facility provided by the school, the service classification requires information on whether facilities are **single-sex** and **usable**.

Single-sex

Single-sex toilets means that separate girls’ and boys’ toilets are available at the school, or the school has only girl or boy students and has toilets. To be considered separate, facilities should provide privacy from students of the opposite sex, but national definitions may be further adapted based on local context. Schools that have separate shifts for girls and boys (i.e. girls attend the school at a different time from boys) are classified as having single-sex toilets as the same toilets are

used separately by girls and by boys. A gender-neutral room with a single private toilet may also be considered single-sex, as it allows individuals to use the toilet privately and separately. The criterion of single-sex is not included in calculations of basic sanitation service in pre-primary schools.

Usable

For the purposes of global monitoring of school sanitation, “usable” refers to toilets or latrines that are (1) available to students (doors are unlocked or a key is available at all times), (2) functional (the toilet is not broken, the toilet hole is not blocked, and water is available for flush/pour-flush toilets), and (3) private (there are closable doors that lock from the inside¹⁵ and no large gaps in the external structure) at the time of the questionnaire or survey. If data are reported as “functional latrines” these are classified as usable, not improved and usable, while “functional toilets” or “functional bathrooms” are classified as improved and usable in the absence of more detailed data.

Where data on all three of these criteria are available, only schools with toilets or latrines that meet all three elements are considered to have usable sanitation. In cases where data are only available on one or two of these elements, those data are used for estimates in the absence of more detailed data from another source. For example, data on functionality alone is used as a proxy of usability in the absence of information on availability and privacy until more comprehensive data are available.

Calculation of basic sanitation service

The proportion of schools with basic sanitation services is calculated as the proportion of schools with improved sanitation facilities that are usable and single-sex (**Figure 3**). Schools with improved sanitation facilities that are either not single-sex or not usable are classified as having limited services. The remaining proportion of schools with unimproved or no sanitation facilities are classified as having no sanitation service. In cases where there are data on any facility but no data on improved facilities from any dataset, no service is estimated based on the proportion of schools with no facility.

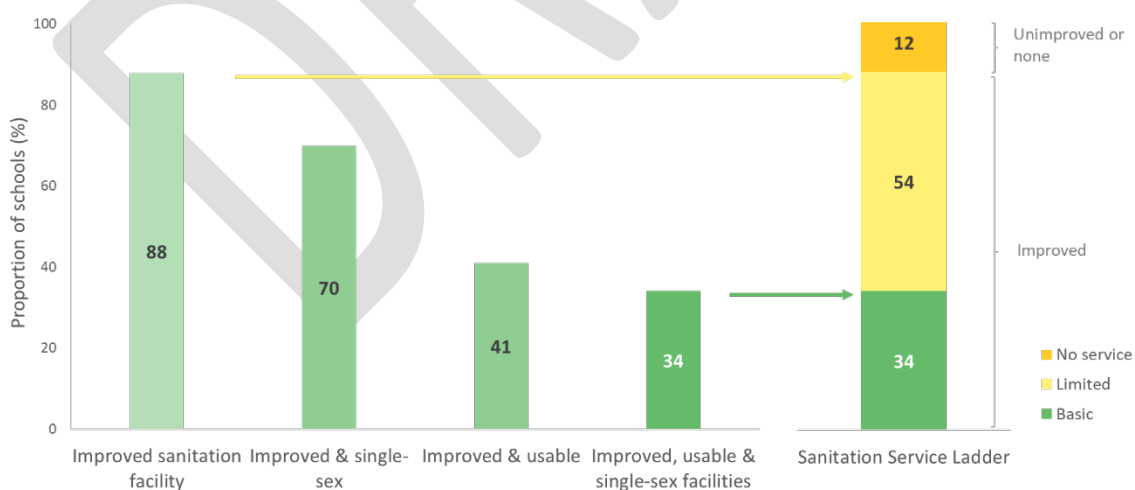


Figure 3. Example of calculating basic sanitation service in schools

In cases where microdata are unavailable, but the proportion of schools with improved sanitation facilities is known and the proportion of schools with usable and single-sex facilities is known, basic is calculated by multiplying the proportion of schools with improved facilities by the ratios of schools

¹⁵ Toilets that are lockable from the inside may not be applicable in pre-primary schools.

with usable facilities to schools with any sanitation facility and of schools with single-sex facilities to schools with any sanitation facility. For example, if 50% of schools have usable toilets, 80% have single-sex, 65% have improved sanitation facilities, and 10% have no sanitation facilities, it is estimated that 32.1% of schools have basic drinking water service ($50/90 \times 80/90 \times 65 = 32.1$).

If data are available on the usability of toilets but data are missing for either improved or single-sex, an estimate of basic service is made based on only the proportion of schools with usable toilets and either the proportion with improved, or the proportion with single-sex facilities. This assumes that usability is the most important criteria. Where the number of toilets is known, but not if they are single-sex, the proportion of schools with basic sanitation service is calculated based on schools with improved, usable, and at least two toilets. Having two or more toilets is not used as a proxy for single-sex separately, however. Estimates based on incomplete datasets are reviewed as new data become available.

2.3.3 Hygiene

In addition to the availability of handwashing facilities at the school, the **basic** service classification requires information on the availability of **soap** and **water** at the time of the survey or questionnaire.

Availability of soap

Soap should be available at the handwashing facilities. Bar soap, liquid soap, powder detergent and soapy water all count as soap for global monitoring purposes. In some cultures, ash, soil, sand or other materials are used as handwashing agents, but these are less effective than soap and are therefore insufficient for basic hygiene service.

Availability of water

Similar to drinking water availability, where schools report water availability on average rather than the day of the survey, schools with water available at least 50% of the time (e.g. at least 4 hours per 88-hour school day or 3 days per 5-day school week) are classified as having water available for handwashing. In the absence of data on water availability, the functionality of the handwashing facility is used as a proxy where functional facilities are assumed to have water available. If there is information on handwashing facilities and soap, but not water availability, data on drinking water availability is used as a proxy, if available.

The JMP uses data from school censuses (e.g. EMIS) and other school-reported sources in which the head teacher or other school administrator observes the handwashing facility and confirms the presence or absence of soap and water. In the few cases where datasets include both observation by an enumerator and school-reported availability of handwashing facilities with soap and water, school-reported data are used to make estimates for consistency with national EMIS which typically only include school-reported data.

Calculation of basic hygiene service

The proportion of schools with basic hygiene services is calculated as the proportion of schools with handwashing facilities that have soap and water available (**Figure 4**). Schools with handwashing facilities with water available but not soap are classified as having limited service. The remaining proportion of schools with no water available or no handwashing facilities are classified as having no hygiene service. In cases where there are data on availability of facilities but no data on facilities with water from any dataset, no service is estimated based on the proportion of schools with no

handwashing facility. Estimates based on incomplete datasets are reviewed as new data become available.

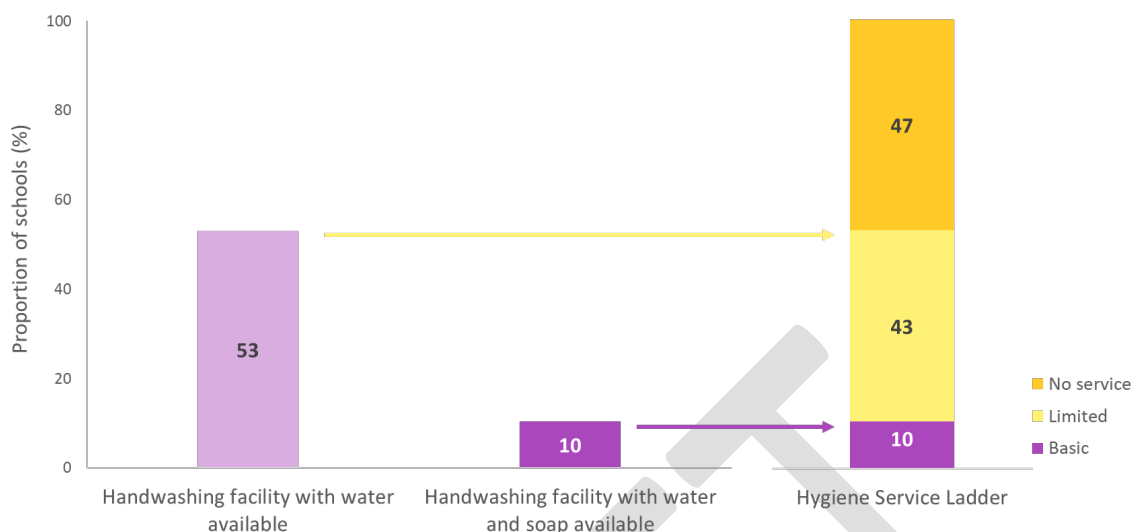


Figure 4. Example of calculating basic hygiene service in schools

2.4 Data cleaning and adjustment

Where microdata are available, the JMP calculates the relevant estimates and JMP tabulations are compared against any tabulations reported in, for example, survey reports or administrative reports. In the case of inconsistencies, the questionnaire categories are checked to identify possible errors.

Where response categories are ambiguous or do not map clearly into JMP standardized categories, adjustments are made drawing upon data from other datasets for the same country and time period, where available. For example, some surveys record data on schools using ambiguous classifications such as “traditional wells”. Some traditional wells, but not all, meet the criteria for *improved* sources, so these schools cannot immediately be classified as using an *improved* or *unimproved* water source. If another dataset is available which distinguishes between unambiguous classifications (e.g. the proportion of all wells which are improved), this ratio can be applied to the survey with the ambiguous classification. If there are multiple datasets from which ratios can be generated, the dataset closest in time to the ambiguous data is used.

In some cases, all datasets within a country use ambiguous classifications. In this case, there should be a specific consultation with the country, and in the absence of other information 50% of the ambiguous technologies are considered as *improved*, and 50% *unimproved*. In cases where response options include many *improved* facility types and few *unimproved* facility types, ambiguous categories, such as ‘other’, may be classified as *unimproved*.

Some datasets only include information on schools with specific types of facilities that correspond with national standards (e.g. piped water). Data on *piped water* can be used to calculate *improved* by assuming that 50% of the complement (non-piped water) are improved. However, this statistic will only be used for estimating *improved* when *piped* is at least 90%.

Some facility type data are based on a multiple response question where the total proportion of schools using each facility type sums to more than 100%. Where microdata are available, the number of facilities per school is reduced to one based on a hierarchy starting with the ‘most

improved' (e.g. piped water, flush toilet) to the 'least improved' (e.g. surface water, bush/field), so that schools with any improved facilities are classified as such, even if the school also has unimproved facilities. If microdata are unavailable, the data are not used unless additional information is available to facilitate data cleaning and adjustment.

If sanitation or hygiene data are reported separately by sex, the minimum, maximum, or average of the two values is used depending on the indicator. The maximum is used for the proportion of schools with any facility (**Table 2**: S_1 , H_1). For example, if 85% of schools have handwashing facilities for boys and 81% have handwashing facilities for girls, H_1 is 85%. For sanitation specifically, the average of the two is used for the proportion of schools with each facility type (S_2), and the minimum is used for the proportion with single-sex or usable facilities to calculate basic service (S_3). For hygiene specifically, the average value is used for water and soap availability.

Missing data are excluded from the analysis if greater than 5%, unless there is information that suggest these data indicate something else, such as *no facility* or *other*, or data from other sources suggest they should be included in the analysis. If less than or equal to 5%, they are classified as *no facility* or *no service*, as appropriate. These assumptions are checked for consistency and reviewed as new datasets become available.

If any of these adjustments are made in the JMP country files, they are highlighted in the notes for that dataset.

2.5 *Data acceptance*

In some cases, data that have been extracted and shown in the JMP country files are considered unsuitable for use in JMP estimates and excluded from further analysis. The most common reasons for excluding datasets include:

- Communication from national authorities that the data are not considered reliable or appropriate for use.
- The national classification of the data is based on few generic categories which are not aligned with JMP standard categories.
- The "other" category is very large.
- Data are not representative of national, urban or rural, or pre-primary, primary or secondary schools.
- Sums of categories which should be mutually exclusive and exhaustive are far from 100%.
- Data are markedly different from multiple other data points from a similar timeframe.
- Data better aligned with the SDG indicator definitions become available and are markedly different.

For any of these reasons, a dataset may be excluded from calculation of any of the indicators reported by the JMP. In some cases, a dataset can be used for one or more but not all indicators because of variable data availability and quality. For example, a school survey might yield reliable data on "improved sanitation" but unreliable data on the usability of those facilities, because of ambiguous question wording or inadequate training of survey teams. When data shown in a country file are not used for estimates, the reason is recorded in the notes for that dataset.

3. Producing country estimates

3.1 National, rural and urban, and school level estimates

All data that have been reviewed and included in the analysis for a country and domain are used to produce JMP estimates. Separate estimates are made for national, urban, rural, pre-primary, primary and secondary schools where data are available. If data are only available for primary schools, these data are used for national estimates since primary schools are often the majority of schools and serve a large proportion of the population. If data are only available for rural (or urban) schools, these data are used for national estimates where the population living in rural (or urban) areas comprise more than 80% of the national population. See section 2.1 for details on weighting data on rural and urban or school level by their respective populations to produce national estimates.

3.2 Primary indicators

The JMP estimates nine primary indicators directly from data inputs, all of which refer to the proportion of schools with access to drinking water, sanitation, or hygiene facilities or services (**Table 2**). Separate estimates are made for each indicator. In many cases data are only available for some of these indicators. For example, some countries may have an estimate for the proportion of schools with improved sanitation facilities (S_2), but no data on the elements of usable and single-sex needed to calculate the proportion of schools with a basic service (S_3). In other cases, data are reported for the proportion of schools with basic sanitation (S_3), but separate figures are not provided (and microdata are unavailable) for the proportion with improved facilities (S_2).

For each primary indicator with data available, estimates are calculated using a standard set of “JMP estimation” rules for interpolation, extrapolation, and extension of estimates.

Interpolation rules (within the years covered by data inputs)

1. If data points are at least four years apart, estimates are interpolated using ordinary least squares linear regression¹⁶
2. If data points are less than four years apart, estimates are interpolated using a simple average

Extrapolation rules (outside the years covered by data inputs)

1. If there are at least two data points, estimates are extrapolated using ordinary least squares linear regression (or simple average, if data points are less than four years apart) forwards two years from the latest data point, and backwards two years from the earliest data point
2. If extrapolation results in estimates above 99% or below 1%, the estimates are fixed at 100% and 0% and reported as “>99%” and “<1%”, respectively
3. If there is only one data point, no extrapolation is made

Extension rules (beyond the years covered by data inputs and extrapolation)

1. Estimates are extended forwards four years from the latest estimate, and backwards four years from the earliest estimate

¹⁶ This follows a similar method to that used for household data, which uses a span of five years. A shorter time span is used for school data based on the common usage of annual EMIS data (or annual data from the same source).

- Estimates are extended forwards or backwards indefinitely, if the estimate is 99.5% or greater, or 0.5% or less.

Estimates are made for all years possible within the period from 2000 to one year prior to the year of reporting¹⁷. The following examples illustrate the application of JMP estimation rules for the 2020 progress update which included estimates for all years between 2000 and 2019.

Examples

Figures 5-9 provide illustrative examples of applying the rules for interpolation, extrapolation, and extension of estimates.

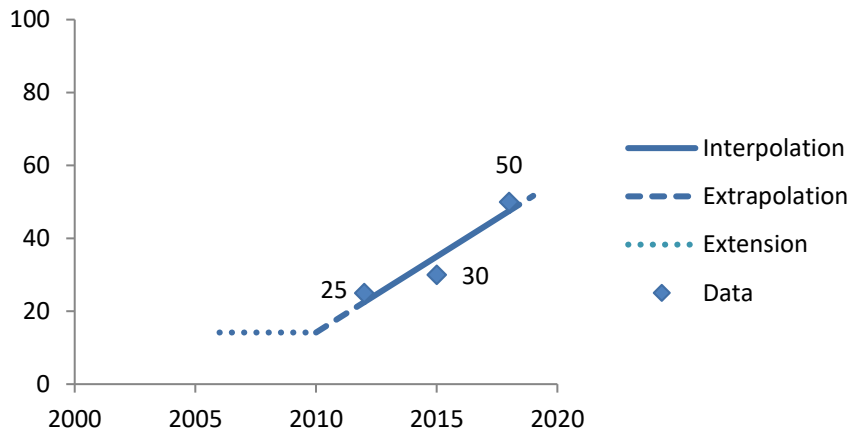


Figure 5. Three data points are available, spanning at least four years. Interpolation is made from 2012-2018 with linear regression. Extrapolation is made backwards for two years, from 2012 to 2010, and this estimate is extended backwards for four years, until 2006. Extrapolation is made forwards from 2018 to 2019, which is the end of the reference period for the JMP 2020 progress update (2000-2019).

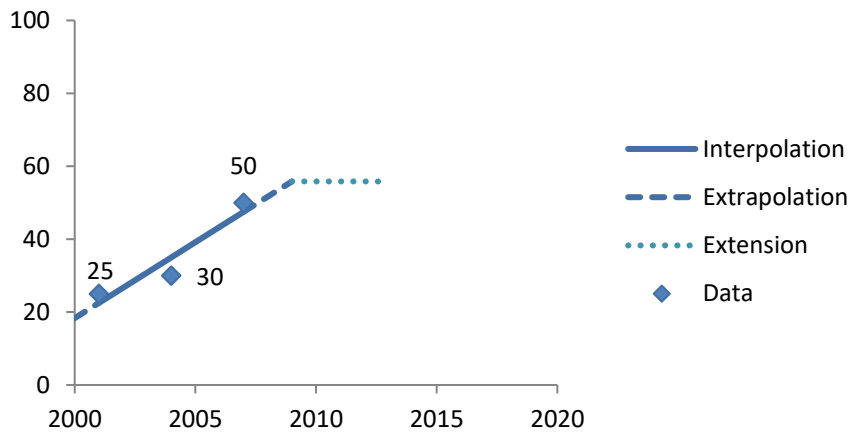


Figure 6. Three data points are available, spanning at least four years. Interpolation is made from 2001-2007 with linear regression. Extrapolation is made backwards from 2001 to 2000 which is the beginning of the reference period. Extrapolation is also made forwards from 2007 to 2009. The 2009 estimate is extended for four years to 2013, and no estimate is made for 2014 and beyond.

¹⁷ In the 2018 baseline report estimates were made for all years possible within the period from 2000 to two years prior to the year of reporting (i.e. 2000 to 2016).

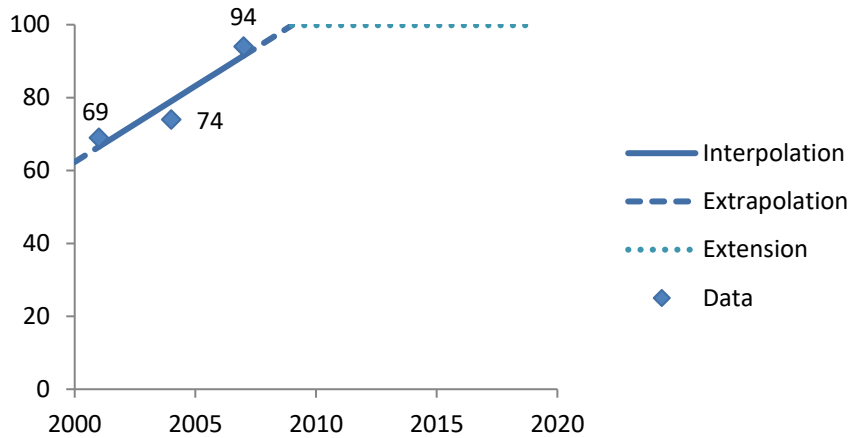


Figure 7. Three data points are available, spanning at least four years. Interpolation is made from 2001-2007 with linear regression. Extrapolation is made backwards from 2001 to 2000, and forwards from 2007 to 2009. The 2009 estimate of 99.8 % is extended indefinitely since it is above 99.5% to 2019 which is the end of the reference period.

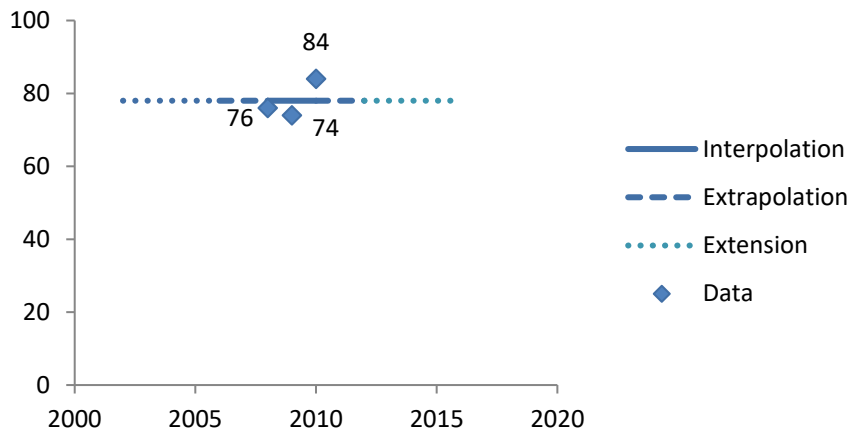


Figure 8. Three data points are available, spanning less than four years. Interpolation is made from 2008-2010 by averaging. Extrapolation of this averaged value is made backwards from 2008 to 2006, and forward from 2010 to 2012. The 2006 estimate is extended backwards to 2002, and the 2012 estimate is extended forward to 2016. No estimate is made for 2017 or later.

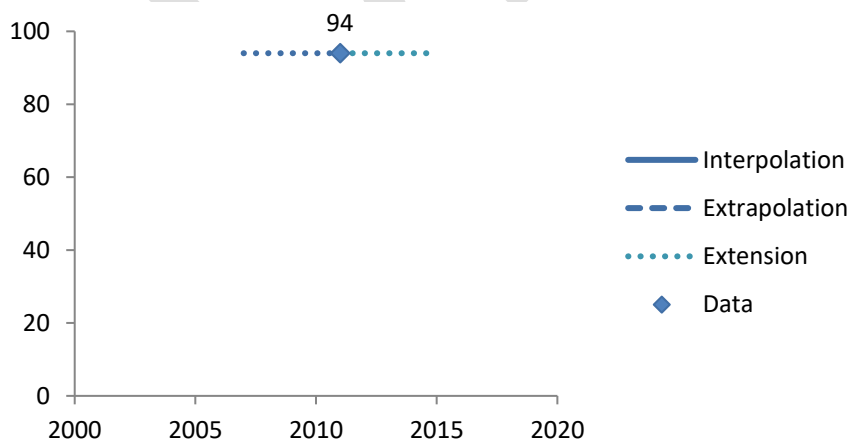


Figure 9. Only one data point is available. No interpolation or extrapolation is made. The 2011 estimate is extended backwards to 2007, and forward to 2015. No estimate is made for earlier or later years.

Box 1. Evaluation of alternatives to linear regression

In 2015 the JMP convened an Expert Group Meeting to review the methods used for interpolation, extrapolation and extension. Alternatives to linear regression were considered including logistic regression, piecewise spline, and loess. While the datasets in the JMP database show evidence of some nonlinear trends for some parameters over time, the nonlinear models assessed require large numbers of datapoints to allow interpolation of estimates that are superior to simple linear regression. For many countries, there are not enough datapoints to justify the use of nonlinear methods. Nonlinear techniques are also limited in their ability to extrapolate even a few years, which is often necessary for the JMP's estimation method. Accordingly, the JMP has continued to use linear regression, with the set of rules regarding interpolation, extrapolation and extension described above¹⁸.

3.2.1 Adjustments

Improved drinking water or sanitation facilities is a subset of any facilities, so data on any facility (W_1 and S_1) should never be greater than improved water (W_2) or improved sanitation (S_2), respectively. Similarly, basic service is a subset of improved facilities, so data on improved facilities (W_2 and S_2) should never be greater than basic drinking water service (W_3) or basic sanitation service (S_3), respectively. Because separate regressions are made for each of these indicators, and may draw on different sets of data, it is possible for these unrealistic situations to be estimated, in which case the JMP applies standard corrections. If the estimate for improved facilities exceeds the estimate for any facility, W_2 is set as equal to W_1 , unless datapoints for improved are considered more reliable, in which case W_1 is set as equal to W_2 . If the estimate for basic service exceeds improved facilities, W_3 is set as equal to W_2 . The same logic is used for sanitation. For hygiene, if the estimate for handwashing facility with water available is higher than handwashing facility, H_2 is set as equal to H_1 , unless datapoints for facility with water available are considered more reliable, in which case H_1 is set as equal to H_2 . If the estimate for basic hygiene service exceeds the estimate for handwashing facilities with water, H_3 is set as equal to H_2 .

JMP estimation rules may result in estimates that predate the establishment of a new country, area, or territory. Estimates are removed in such cases:

- No estimates are made for South Sudan before 2011
- No estimates are made for Montenegro before 2006
- No estimates are made for Timor-Leste before 2002

3.3 Annual Rates of Change

In order to calculate an annual rate of change (ARC) for different indicators, the difference between two estimates from different years is divided by the number of years separating the estimates:

$$ARC_x = \frac{P_{x,y2} - P_{x,y1}}{y2 - y1}$$

¹⁸ For further information see <<https://washdata.org/report/jmp-methods-task-force-report-final>>

Where ARC_x is the annual rate of change for indicator x

$P_{x,y1}$ is the estimate for indicator x in reference year $y1$, in percentage points

$P_{x,y2}$ is the estimate for indicator x in reference year $y2$, in percentage points

For example, the 2020 JMP progress update on WASH in schools calculated ARCs for three variables (basic drinking water services, basic sanitation services, and basic hygiene services) using 2015 and 2019 as reference years. Where national estimates were lacking for either reference year, no ARC was calculated. A negative ARC represents a declining proportion of schools with basic services.

4 Country consultation

Every two years the JMP updates its global databases to incorporate the latest available national data for the global SDG indicators. The JMP is committed to consulting national authorities on the estimates generated from national data sources through a country consultation process facilitated by WHO and UNICEF country offices. While the two agencies work together, generally UNICEF leads the consultation efforts for WASH in schools. The country consultation aims to engage national statistical offices and other relevant national stakeholders to review the draft estimates and provide technical feedback to the JMP team.

The purpose of the consultation is not to compare JMP and national estimates of WASH coverage but rather to review the completeness or correctness of the datasets in the JMP country file and to verify the interpretation of national data in the JMP estimates.

The JMP produces detailed guidance to facilitate country consultation on the estimates contained in JMP country files. A key resource is the Data Summary tab of the country file which lists the national sources of data which have been collected to date (see Annex 1). The consultation focuses on three main questions¹⁹:

- a. Is the country file missing any relevant national sources of data on drinking water, sanitation and hygiene in schools that would allow for a better estimate?
- b. Are the data sources listed considered reliable and suitable for use as official national statistics?
- c. Is the JMP interpretation and classification of the data extracted from national sources accurate and appropriate?

For the purpose of global monitoring and comparison the JMP seeks to apply the same rules to all countries and for this reason some common requests from countries cannot be accommodated, including:

- To use national indicators instead of SDG global indicators for WASH in schools
- To use national data on rural, urban and school age populations, rather than standardised estimates produced by UN agencies
- To assume that 100% of schools have basic WASH services, if there are regulations or standards requiring basic (or higher) levels of service, without supporting data to verify compliance rates

¹⁹ The latest full country consultation guidance note can be found at <https://washdata.org/report/jmp-2020-schools-country-consultation>. It is also available in [Spanish](#) and [French](#).

- To make provisional estimates based on partial datasets or small studies which did not meet the minimum threshold for data coverage required under the JMP rules described above
- To simply use the most recent data point, rather than producing modelled estimates based on all available data points within the reference period

WHO and UNICEF endeavour to consult with all countries and respond to feedback received, and to give explanations in cases where JMP definitions or methods differ from those used or recommended by national stakeholders.

5 Producing regional and global estimates

5.1 School age population

In the absence of a global database with consolidated and consistent data on the number of schools in each country, the JMP uses school age population data to weight the individual country estimates when producing the regional and global estimates. Pre-primary, primary and secondary school age populations are extracted from the most recent data published by the UNESCO Institute of Statistics (UIS)²⁰, which may differ from national statistics.

For the purposes of generating regional estimates, the national school age population is calculated as the sum of the pre-primary, primary and secondary (including lower secondary and upper secondary) school age populations. Urban and rural school age populations are calculated using the percentage of population residing in urban areas as reported by the UN Population Division.

School age population data are available for most countries but remain incomplete. For example, in 2019, 186 countries, areas and territories had primary school age population estimates (out of a total of 234 for which we have population estimates from UN Population Division). The JMP uses two approaches for the imputation of missing data for school age populations:

1. **Countries with incomplete time series:** The school age population for countries which do not have data for all years within the reference period is imputed using linear regression (**Figure 10**). These extrapolated (or interpolated) estimates are included in grey in the JMP country files for schools and flagged as “based on linear regression”.
2. **Countries with no data:** Imputation for countries with no school age population data is based on the ratio of school age children to the total population in each M49 subregion²¹. These imputed values are included in grey italics in the JMP country files for schools and flagged with a note: “In the absence of UIS data on school age population, values in grey are imputed based on the M49 regional mean for school age population”.

The JMP uses UIS regional and global estimates of school age population and multiplies these by the proportions of schools with different WASH service levels to estimate the regional and global school age population with each service.

²⁰ “School age population by level of education” data series in the UNESCO Institute for Statistics database (UIS.Stat): <http://data.uis.unesco.org/> (under “Education” theme and “National Monitoring”-“Population of the official age / school age population domain”).

²¹ M49 stands for the UN Standard Country or Area Codes for Statistical Use: See United Nations Department of Economic and Social Affairs, *Methodology* for more details: <<https://unstats.un.org/unsd/methodology/m49/overview>>

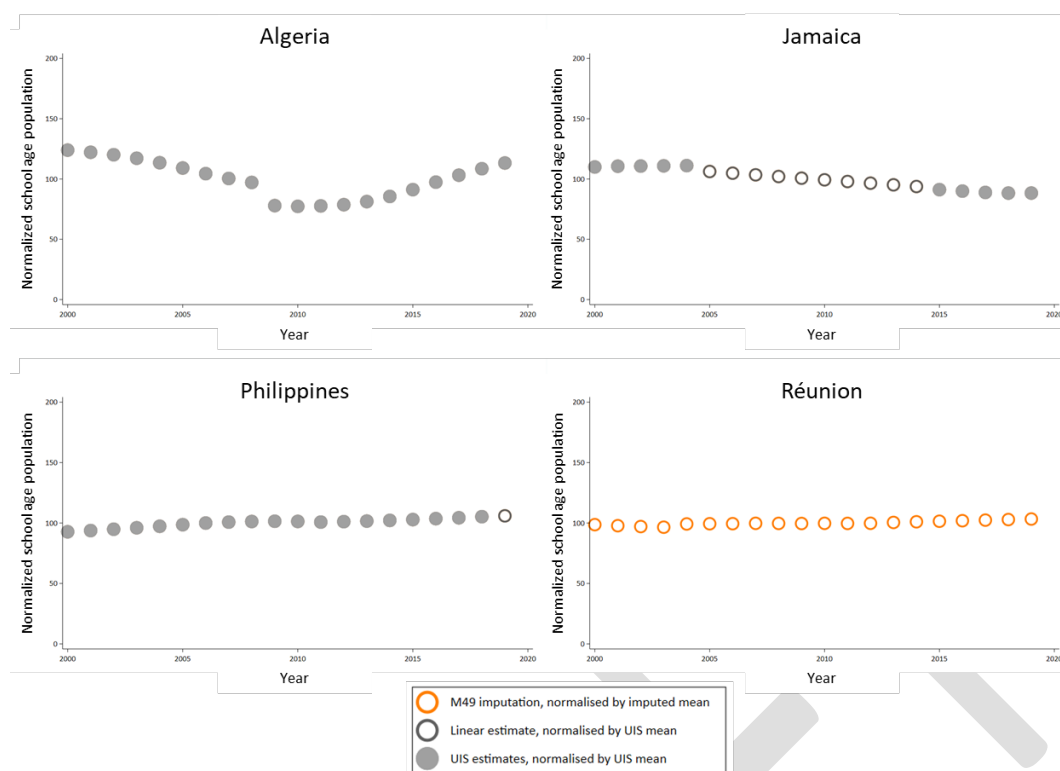


Figure 10. Examples of imputations for primary school age children: Algeria (no imputation required), Jamaica (interpolation), Philippines (extrapolation), Réunion (no data). All charts above are normalized by the average school age population 2000-2019.

5.2 Service level estimates

Regional estimates for basic drinking water, sanitation and hygiene services are calculated using simple school age population weighted averages for each country, area or territory in the region. Aggregated estimates are only made when the school age population with data coverage exceeds 30%.²²

Global estimates are made in a two-step process by aggregating school age population weighted regional estimates rather than averaging weighted country estimates directly. In order to produce global estimates, national estimates are imputed for countries lacking data for one or more of the WASH in schools indicators (**Table 2**). Imputed national estimates are not published and only used for global aggregation.

Imputations are made by calculating the school age population weighted average for the relevant “master region”. The SDG regional classification is used as the master regional grouping for values published in the JMP progress updates to date (**Table 5**).

Table 6. SDG regional groupings used for the 2020 update report.

SDG region	Number of countries, areas and territories	School age population in 2019 (thousands)
Australia and New Zealand	2	5,367
Central and Southern Asia	14	549,883
Eastern and South-Eastern Asia	18	448.945

²² A cutoff of 50% is used for basic WASH services in households. A lower threshold is temporarily set for schools due to data scarcity for these new global indicators.

Europe and Northern America	53	186,470
Latin America and the Caribbean	50	152,102
Northern Africa and Western Asia	25	136,350
Oceania	21	4,504
Sub-Saharan Africa	51	398,484
Total	234	1,882,111

Estimates are reported for multiple regional groupings. The 2020 update report included estimates for the global region, the SDG regions, and three thematic groupings: Least Developed Countries (LDCs), Landlocked Developing Countries (LLDCs), and Small Island Developing States (SIDSs). The same methodology is used to calculate regional estimates for other regional groupings. Estimates for the following regional groupings are currently available on the JMP website²³: World, SDG regions, LDCs, LLDCs, SIDSs, M49 regions, M49 sub-regions, WHO, UNICEF programme, UNICEF reporting, UNESCO GEMR, UNESCO UIS and World Bank income groups.

Box 2. Potential sources of uncertainty

There are multiple steps in the process of calculating national, regional and global estimates for WASH in schools, as described in this document. There could be a large degree of uncertainty in resulting estimates due to the uncertainty that is associated with each step of this process, including:

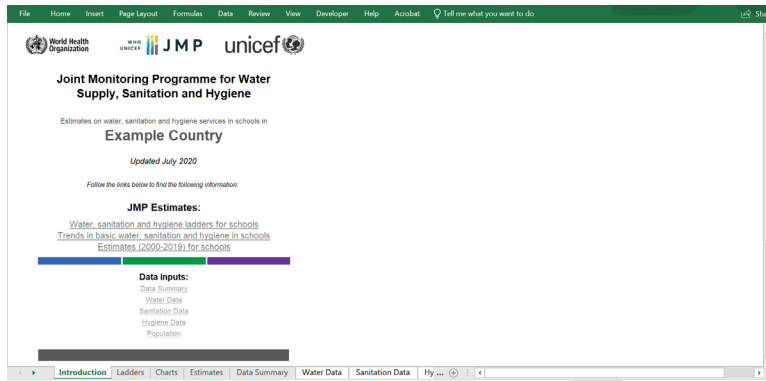
- **Individual data sources** may have sampling error due to sample size, non-sampling error due to missing populations such as the exclusion of private schools from some EMIS, and ambiguous terms that are not easily comparable to other data sources.
- **National estimates** can have modelling errors around regression (see Box 1).
- **Regional estimates** may have errors due to low data coverage within the region.
- **Global estimates** may have errors due to imputation of estimates for countries with missing data.

Uncertainty is decreasing with each progress update as countries are harmonizing data collection activities, reducing ambiguity and filling data gaps. These improvements are reflected with new regression lines each time data is added. Each new progress update therefore supersedes the last.

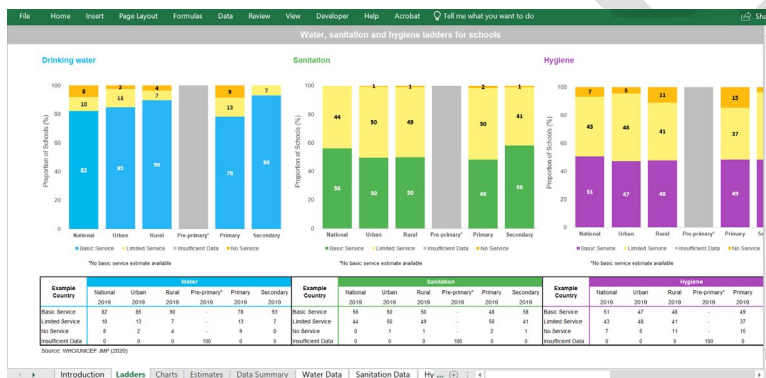
²³ www.washdata.org

Annex 1: JMP Country Files for WASH in schools

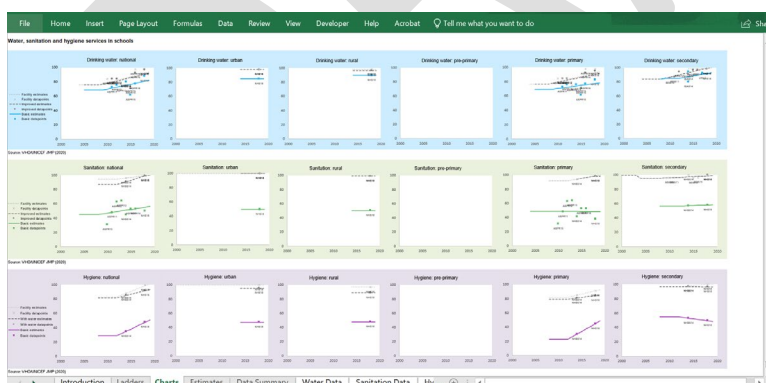
JMP country files have been created in Excel for drinking water, sanitation and hygiene in schools based on the scope and ambition of the SDG targets. The country files detail the data sources that are available in the JMP database, as well as the resulting JMP estimates. The following outlines each tab of the country files for WASH in schools. Country files are available for each country with data on the JMP website.²⁴



Introduction: The Excel spreadsheet has a series of tabs but for ease of reference the front page includes links to the key tabs.



Ladders: This tab displays drinking water, sanitation and hygiene 'ladders' used by the JMP for global monitoring purposes. The ladders show the latest national, urban, rural, pre-primary, primary and secondary school estimates. Summary estimates are tabulated below as they will appear in the statistical tables at the back of JMP progress reports on schools.



Charts: This tab includes charts showing estimated coverage trends for drinking water, sanitation, and hygiene services since the year 2000. These charts illustrate the JMP method of using a linear regression of available data points to generate estimates for a given reference year and highlight differences between estimates of 'improved,' which has been a common indicator used in the past, and estimates of 'basic' services.

²⁴ <https://washdata.org/data/downloads#WLD>

Annex 2: Data flows in the JMP country file for WASH in Schools

As described in Annex 1, the Excel country files contain sheets for recording the data inputs available for drinking water, sanitation, hygiene, and population. These are connected to two additional sheets – the “Data Summary” sheet and a hidden sheet “Chart Data”. “Chart Data” is a copy of the “Data Summary” sheet with unused (bracketed) data removed.

A statistical analysis software package (Stata) imports data from the “Population” sheet and the “Chart Data” sheet for all the countries/areas/territories where data are available and runs the estimation model. The resulting estimates are exported back to each country file in a hidden “Regression” sheet, which in turn feeds the “Estimates” tab where estimates from 2000 to the reference year are produced for the full set of indicators. The “Estimates” tab, together with the hidden “Chart data” tab, is used to create all the charts and the ladders at country level.

