Problems in representative sampling in the water and sanitation sector: a brief analysis of problems and possible solutions

What is sampling?

Sampling is part of statistics, the scientific discipline and tool used for investigation of biological and medical science in particular. Sampling is the tool used to select part of a population for data collection and analysis. This selection, the sample, is then used as a manageable number of people or objects (depending on what is being investigated) to then form the basis for analysis. In many cases, collecting data for the entire target population would be too expensive in terms of time and resources, as well as too challenging logistically.

We use sampling in every day life. For example, if buying a large quantity of fruit or vegetables from a market, we would probably not check the quality of every item, but instead might check a few for damage and disease. From that we would then make a decision about whether we were happy with the quality of them all.

Sampling techniques can be borrowed and applied by many people in their work. For instance, the accuracy and credibility of field surveys in the water supply and sanitation sector can be enhanced by following accepted sampling techniques.

Why would you use sampling in the water and sanitation sector?

Sampling is used in the sector for a number of different reasons.

1. Monitoring

Monitoring in the sector is carried out by a large number of different actors and for a variety of reasons. At national level, large-scale household surveys are carried out to provide data representative of the whole nation. For instance, OCR-Macro carries out the Demographic and Health Survey (DHS) and UNICEF is responsible for the Multiple Indicator Cluster Survey (MICS). These national data sets provide information about both water supply and sanitation coverage which is used by the WHO/UNICEF Joint Monitoring Programme in monitoring world progress towards the Millennium Development Goals.

id satios usm Dd sampari+h.2 (a)-1.4 I dhee ie
their needs. This may involve field staff selecting households from the whole community and using a questionnaire to collect information from each of the selected households. Similarly, aid agencies might use field surveys to inform and evaluate their water supply and sanitation programmes.

information were required about different social or ethnic groups or urban and rural populations. To carry out stratified sampling, sufficient information is usually required prior to the data collection.

Most common sampling methods will aim to have a self-weighted sample. This means that no sample weights are required. Weighting of data is used to correct imbalances in the probabilities of selection. For instance, if village A has twice as many respondents in the sample than village B (but A has a much smaller population size than B) then weighting can be used to weight up the under-represented village B respondents and weight down the over-represented village A respondents. Weighting is important as it has an impact on the results but is best done by statisticians using appropriate software.

As seen by the challenges mentioned above, the water and sanitation sector is in need of more suitable sampling methods. These methods will need to be able to deal with situations in which there is a high degree of clustering (often making the construction of sample frames impossible (Bostoen, in print)). These methods could also be designed to be more user-friendly for pedly me(n)59ro b en make