

EPA Standard Water Data Terminology and Standard Data File Formats Specifications for Data from External Databases

To analyse different datasets together, the data must be in a form where it can be aggregated easily. This requires data to be collected, reported and stored in a standardised format.

A shared standardised terminology and reporting format ensure that:

- All data from each sampling event can be linked.
- Data collected from different sampling events can be directly compared and analysed.
- Data collected under different sampling programs can be similarly compared and analysed.

Having a prescribed standard reporting format and terminology enables the EPA to progressively aggregate baseline water quality data sets to improve its assessment, regulatory and advisory roles.

This document defines EPA standard terms and file format with respect to water data reporting.

Header Lines

The first two lines of the files are two headers lines; the first line contains the **Report_Code**, which is the code under which the data is submitted; usually, this is the same as the **Program**, but if there are multiple programs submitted in the same file a different code will be used for the **Report_Code**. The 2nd header line is a general description of the data, e.g. *Monthly Sampling May 2022*.

Column Headings Line

The 3rd line is the column headings. The column headings for each field are listed below in the file format section in the **Field** Column. Again, it is crucial the name for each field is the same as the name given for that column in the table in the File Format section below.

File Format

Field	Definition
Site_Code	The Site Code given to the monitoring location by the EPA in the Monitoring Schedule.
SiteName	The name of a monitoring site.
LocName	The name of the sample point that is associated with a site.
Well_Code	For Ground Water Only.
Dry	Set to true if no water present during site visits. If the field is blank, assumed to be false. Can also be used for surface water sites that are dry, or ground water bores with no water present.
Program	The Program to which the sample belongs.
Org_SampleEvent_ID	The organisation's unique Sample Event ID for the Sample Event (if available).
Sample_Date	The Date of the Sample Event. The Date should be the same for all measurements taken at the same Sample Event in the Australian Date Format DD/MM/YY.

Field	Definition
Sample_Time	The time of the Sample Event, in an AEST format. The Sample_time is often the time of the first measurement of the Sample Event. Note: This field is used to identify a sample event data basket. The actual time of the measurement or sample collection is recorded in the 'Depth_Time' field.
TakenBy	The name of the person who took the measurement.
SampleComments	Any comments about the sample.
Sample_Action	The 'Sample Action' taken to obtain the measurement. i.e. LAB, FLD, DUP. Quality Assurance measurements: TRIP, FLDB (Field Blanks), TRIPB (Trip Blanks).
Depth_m	Depth in metres of where the sample was collected in the water column. If not provided depth is assumed to be 0m. Counting from surface as 0 depth, and is to be a positive figure (>0).
Depth_Time	The time the measurement was taken at a particular depth.
Depth_Comment	Where the rationale for taking samples at particular depths is detailed in the sample plan, that information is to be reported here. The field can be used in Groundwater if the bore is monitored at different Aquifers at different depths in the bore.
Lab	Add Name of Lab. For field tests, add 'FLD'.
Lab_Report_No	The Report No issued by the laboratory for the report that contained the Lab result.
Lab_Sample_No	The Sample No issued by the laboratory for the lab sample that was analysed to produce the Lab result.
Lab_Method_No	The name of the Method the Laboratory used to analyse the result.
Indicator_ID	The Indicator ID is given to the Parameter by EPA in the Monitoring Schedule or can be found in the Indicator Register on the EPA web site.
Filtered	F if filtered, if blank assumed to be not filtered.
Parameter	Parameter measured.
Unit	The unit of the measurement.
ResQual	Required for the interpretation of the reported value: > : measurement is greater than reported value. < : measurement is less than reported value. ~ : measurement is an approximate blank : measurement is exact result.
ResVal	The value measured for the Parameter.
Quality	Quality of measured value. 1=Good 2=Average 3=Poor 4=unknown 5=Vetted (EPA internal use only) 6=Hold time issue at lab 7=calculated value

Field	Definition
	If blank assume 4 (unknown)
Sample Method	Niskin Bottle, Pole sampler, integrated depth sampler, Field instrument (make and model eg: YSI EXO3).

Additional Requirements

As well as submitting results of measurements for the required parameters, there are additional requirements when submitting data to the EPA:

- The EPA database uses the concept of a 'Sample Event' to create a basket of data that was collected during a sampling site visit. All measurements taken during the same 'Sample Event' must be linked together by a unique combination of fields, i.e. Site_Code, SiteName, LocName, Sample Date, Sample Time (See Section 'Sample Event').
- During a 'Sample Event', if water quality is measured at different water depths, the results provided must include the water depth in metres at which the laboratory sample and field measurements are taken. Depth must be reported as an attribute of the sample (i.e., in a separate column), not as a measured parameter.
- If 'Depth Labels' are used by the monitoring program to identify samples collected through the water column, the data submitted must include the appropriate Depth Label, i.e. Surface, Middle, Bottom. If there is a requirement to provide 'Depth Labels', the rationale should be detailed in the sample plan.
- Measurements belonging to the same 'Sample Event' can only belong to one Program. In the data file submitted to the EPA, each measurement row belonging to the same 'Sample Event' must have the same Program name in the Program field. (See Section 'File Format').
- For a single 'Sample Event', for each water depth, there can only be one recorded measurement for any specific parameter. Therefore, there should not be in the data file two measurements for the same Parameter with the same water depth for the same 'Sample Event' (even if there are slightly different Sample Times.).

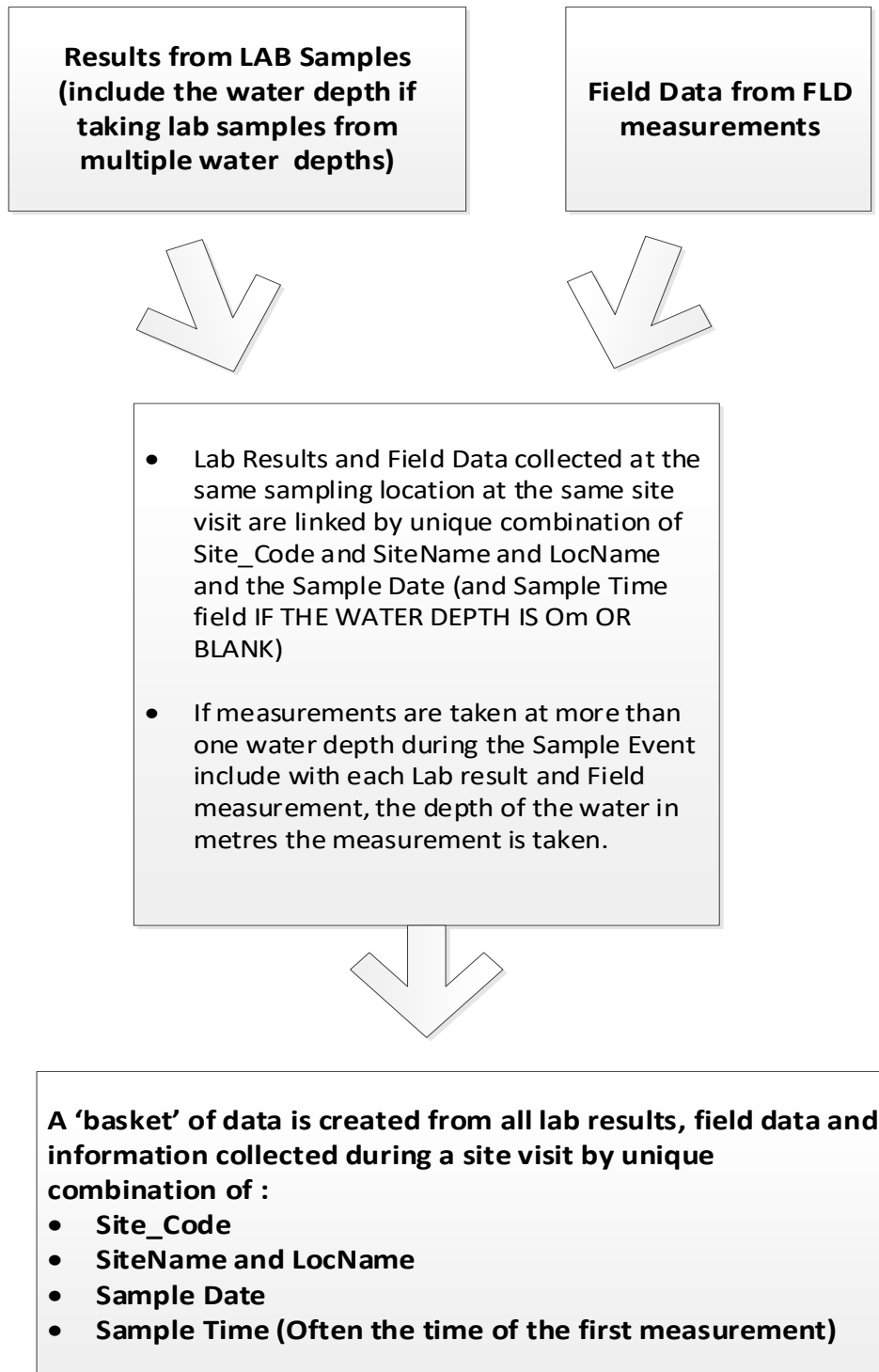
'Sample Event' to create a 'basket' of information

The 'Sample Date' and 'Sample Time' fields provide the key to link all field measurements and laboratory results for a particular sampling visit to a location. The result for each measurement or laboratory result is taken at the same sample location on the same visit is linked by having the following common fields:

- The same Site information ('Site_Code' and 'SiteName' and 'LocName' fields). Where 'Site_Code' is the unique field given to the Site by the EPA.
- The same Sample date ('Sample Date' field).
- The same Sample time ('Sample Time' field). Note: the actual time of the measurement or sample collection is recorded in the 'Depth_Time' field. Often the 'Sample Time' is the time of the first measurement in the Sample Event.

The EPA database expects all samples taken at a location during the same site visit to have a common 'Sample Event' ID (created by the EPA from the above fields) regardless of the depth at which the sample was collected.

Concept of linking Water Quality Measurements



Sample Action

Action taken to obtain measurements. A 'Sample Action' could result in more than one measurement. The total number of 'Sample Actions' for each location makes up a 'Sample Event'. Sample actions include both the field measurement and the laboratory analyses associated with the event, including Quality Assurance analyses (where reported). Possible sample actions and suggested sample action IDs are tabled below:

Standard Sample Actions	Code
Sampling at a monitoring location to be analysed at a Lab	LAB
Integrated sampling (Lund Tube) at a monitoring location to be analysed at a Lab	INT
Using a field instrument to obtain measurements	FLD
QA/QC Sample Actions	
Duplicate Samples	DUP
Triplicate Samples	TRIP
Field Blanks	FLDB
Trip Blanks	TRIPB

The Indicator ID

The Indicator is a chemical, biological or physical property that can be measured. An Indicator can have more than one valid Parameter name because the Parameter is a loose term. To assist in determining the property measured by the Parameter submitted, the EPA has given a unique Indicator for each Parameter required to be monitored in the Monitoring Schedule.

For each Parameter, when providing measurements in an electronic format, the 'Parameter' field provides your system name for that Parameter and in the 'Indicator_ID' field, including the EPA Indicator Code for that Parameter. The Indicator ID for common parameters can be found in the Indicator Register on the EPA website or as provided in the monitoring schedule.

Data File

Water sampling data is to be submitted in a file named "Report_Code.StartDate_EndDate.WaterData.csv".

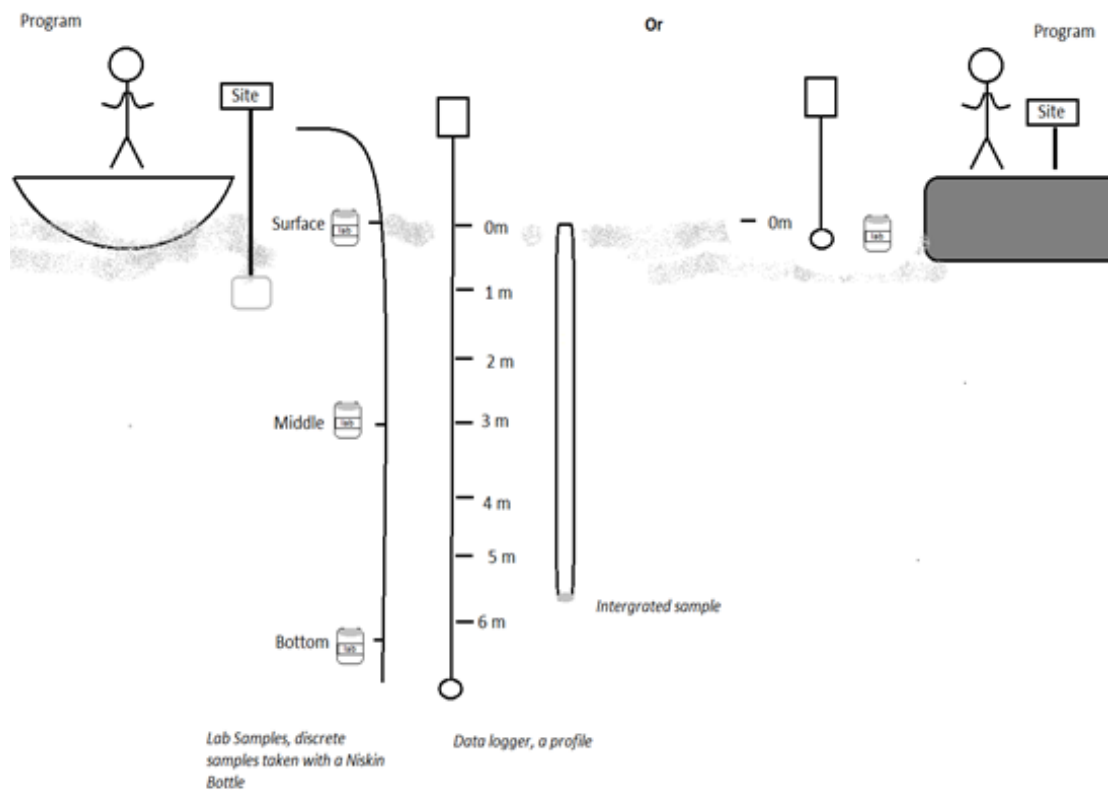
Report_Code: The Report_Code the data is submitted under; often the same as the Program name.

StartDate: The Date of the first measurement.

EndDate: The Date of the last measurement.

Conceptual diagram of sampling

Concept of Sampling



Environment Protection Authority

GPO Box 1550 Hobart 7001

T (03) 6165 4599

E enquiries@epa.tas.gov.au

W www.epa.tas.gov.au



ENVIRONMENT PROTECTION AUTHORITY