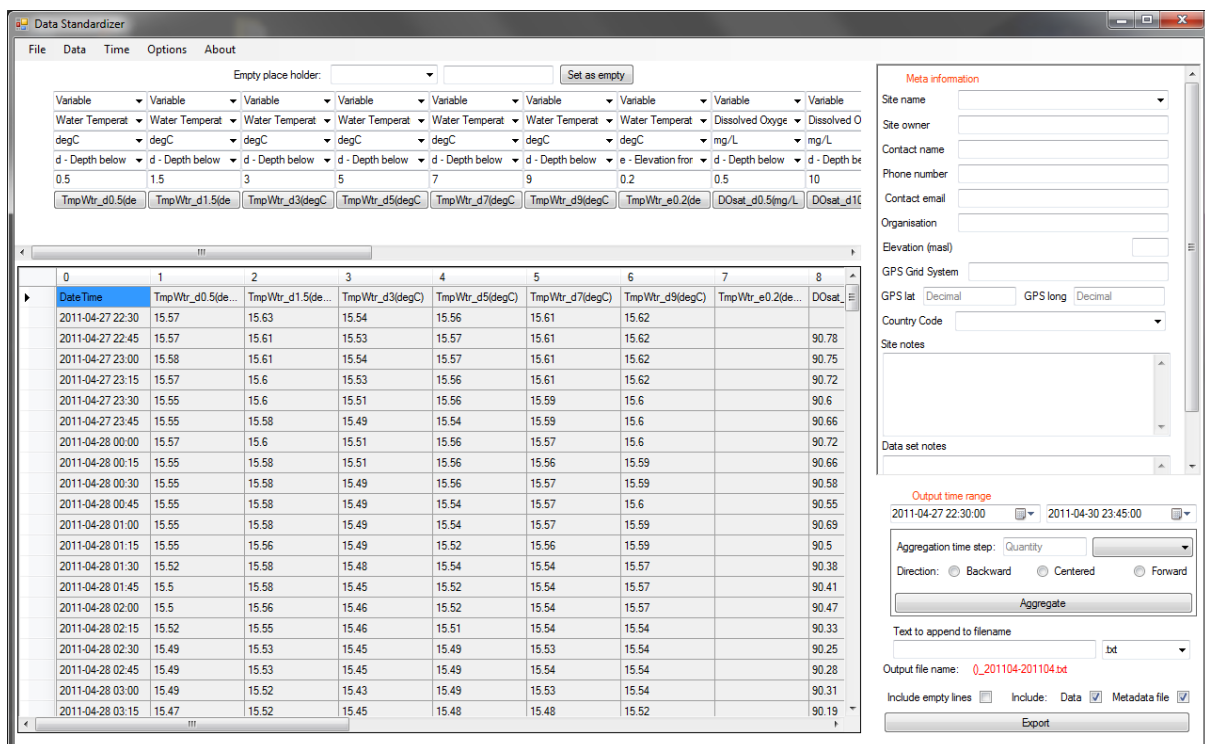


Data Standardizer User's Guide

The Data Standardizer software that is intended to enable easy data importation into B3 as well as Lake Analyzer and other water quality software. It is designed to import large environmental datasets collected at regular time intervals, standardize them and create a separate Meta data file including site information, and then export the date in a variety of formats.

Data Standardizer was developed at the University of Waikato, New Zealand, and was developed by Chris McBride, Richard Lamont and Sam Shute;

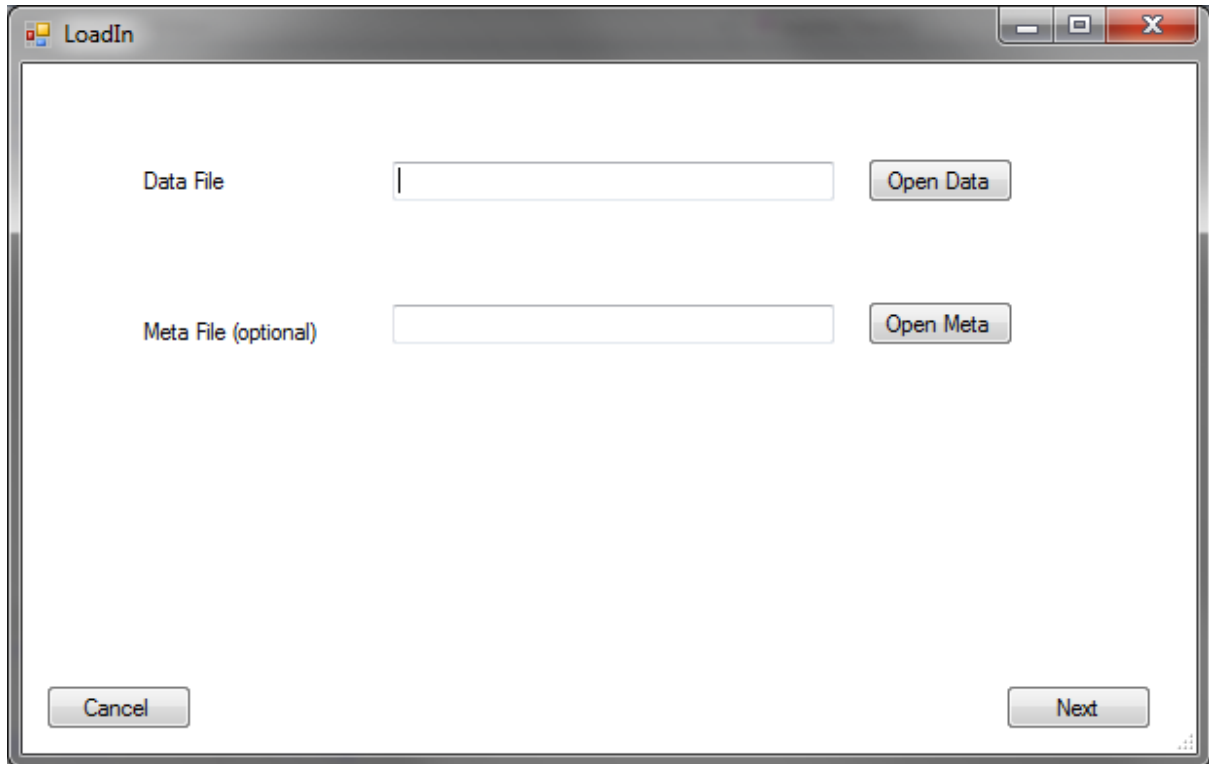
Basic Overview



This is how Data Standardizer will look during normal operation. At the top left is the header bar, with the various drop down header information that you can pick for the corresponding column below. Below that is how your data currently looks, and how it will be exported. To the top right is the segment for Meta Data and below that is the output section, where you can choose how to export your data. You can choose the time range, whether to aggregate the data, the output format and filename.

File Import Process

To begin the file input process, select open dataset from the menu bar under File. You will then be presented with the following screen:



Simply select your data file with the Open Data button, and your Meta file (if you have one from working on this site previously) with the Open Meta button. The Meta file is optional and can be opened later from the Open Meta File from the File menu. You will then be presented with the import options window.

Select Delimiter

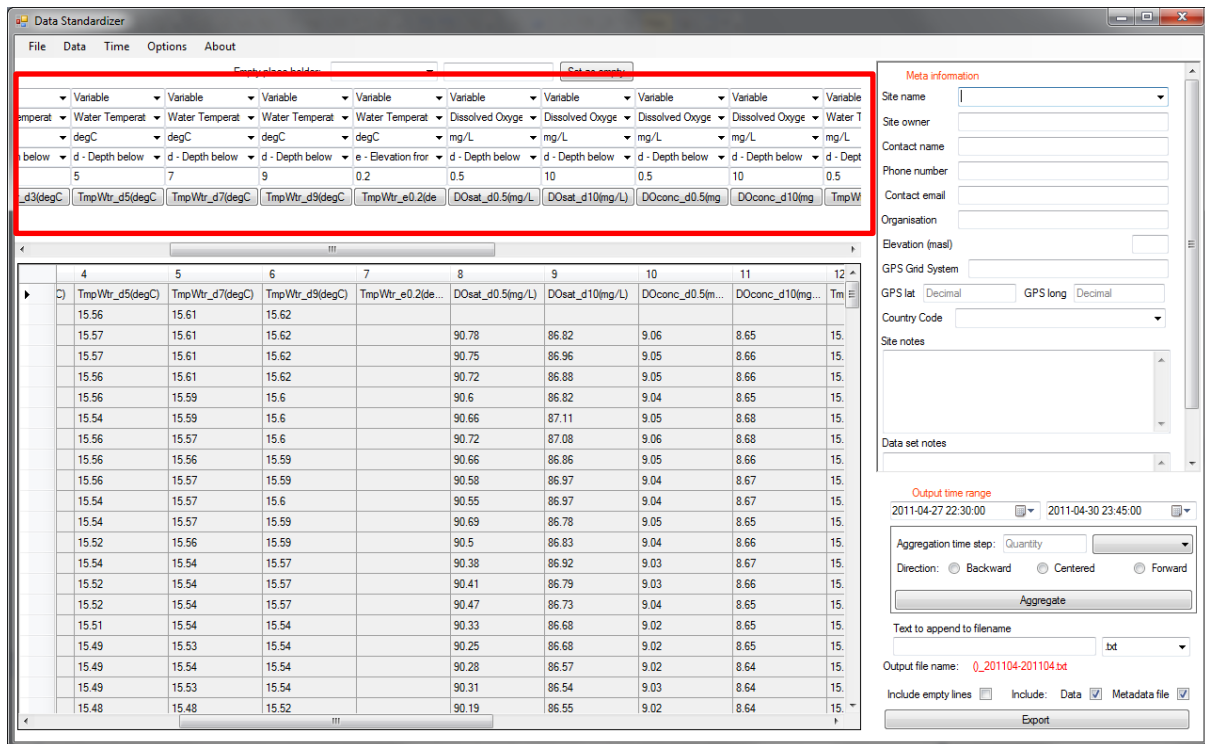
Please Select File Delimiter **OR** Enter a custom file delimiter European Number Format
 Comma Strip incomplete rows shown

	0	1	2	3	4	5
	DD/MM/YYYY	hh:mm	TmpWtr_d0.5(de...	TmpWtr_d1.5(de...	TmpWtr_d3(degC)	TmpWtr_d
▶	27/04/2011	10:30:00 p.m.	15.57	15.63	15.54	15.56
	27/04/2011	10:45:00 p.m.	15.57	15.61	15.53	15.57
	27/04/2011	11:00:00 p.m.	15.58	15.61	15.54	15.57
	27/04/2011	11:15:00 p.m.	15.57	15.6	15.53	15.56
	27/04/2011	11:30:00 p.m.	15.55	15.6	15.51	15.56
	27/04/2011	11:45:00 p.m.	15.55	15.58	15.49	15.54
	28/04/2011	12:00:00 a.m.	15.57	15.6	15.51	15.56
	28/04/2011	12:15:00 a.m.	15.55	15.58	15.51	15.56
	28/04/2011	12:30:00 a.m.	15.55	15.58	15.49	15.56
	28/04/2011	12:45:00 a.m.	15.55	15.58	15.49	15.54
	28/04/2011	01:00:00 a.m.	15.55	15.58	15.49	15.54
	28/04/2011	01:15:00 a.m.	15.55	15.56	15.49	15.52
	28/04/2011	01:30:00 a.m.	15.52	15.58	15.48	15.54
	28/04/2011	01:45:00 a.m.	15.5	15.58	15.45	15.52
	28/04/2011	02:00:00 a.m.	15.5	15.56	15.46	15.52
	28/04/2011	02:15:00 a.m.	15.52	15.55	15.45	15.51

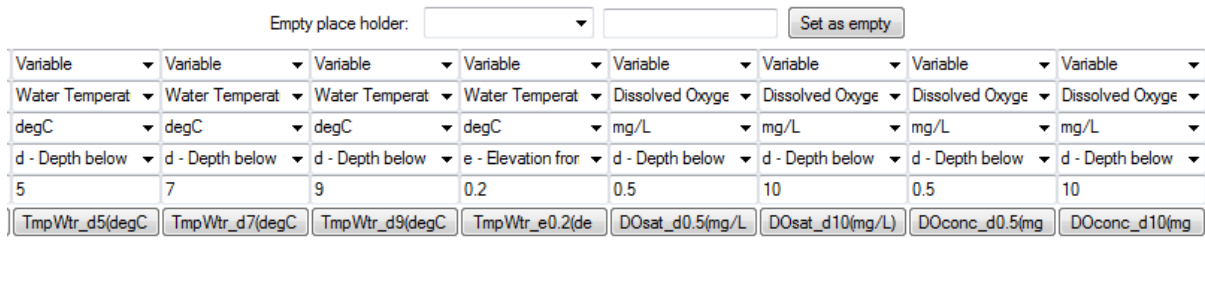
This window allows you to select the header row for your data, as well as an optional Unit Row. Unit rows are when the unit of each column/sensor is in a separate row usually underneath the header row. As well as this you can select if the numbers are in a European format (i.e 1.234.567,89 or 1234567,89), the file delimiter, which can be specified as a custom file delimiter if your file is in an unusual format, and whether to strip the incomplete rows from the sample data.

Once you have completed choosing these settings your data will be imported into Data Standardizer.

Setting the headers



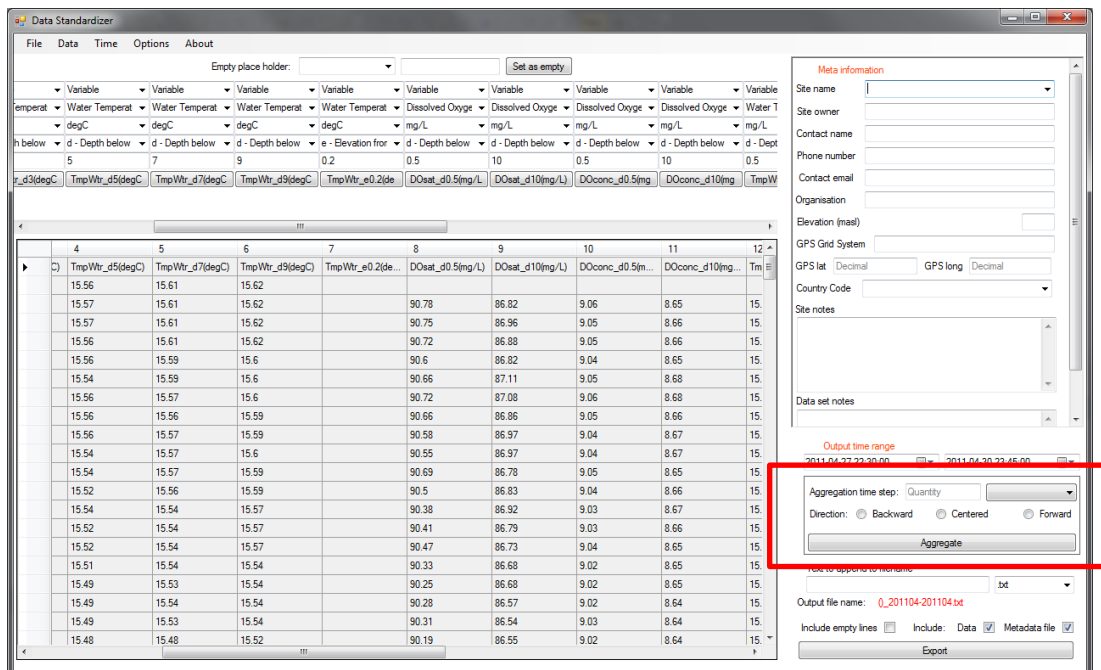
At the top of the screen is the header section. Above each column are they boxes that directly relate to the header for that column. The headers section is designed to set each header up under the Chris McBride method, a standardized format of abbreviations of the gleon control vocabulary.



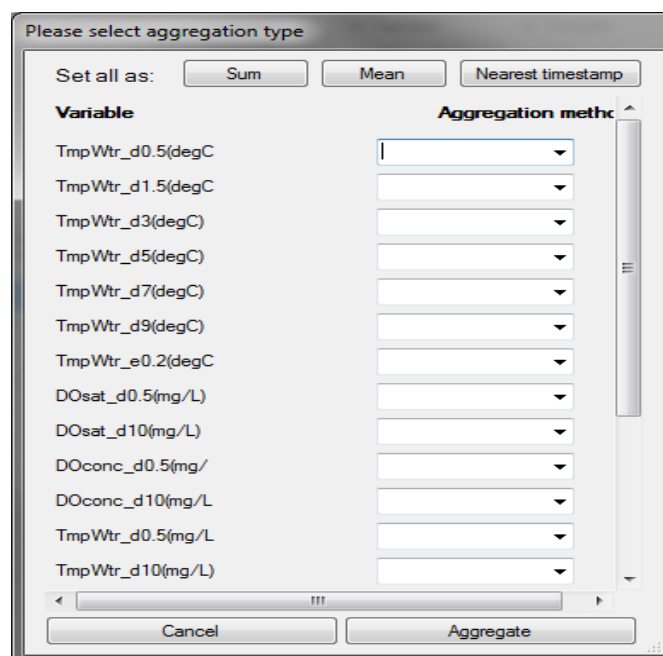
Above is a close up of the actual header boxes. The first sets the type of column, with the options being Variable, Date, Time or the standardized DateTime. Next is sensor type, which is a list of the GLEON control vocabulary. Next is the unit, then sensor positioning and finally the distance measurement for the sensor positioning. For more information on this please see the glossary and the end of this file.

Once you have set up the information in these fields, and are happy with the result on the button at the bottom, you can set the header below by clicking on the button. Or you can set all header from the buttons by going to the Data file menu and clicking Set all from buttons.

Aggregation



One of the most powerful features of Data standardizer is its ability to aggregate data. The aggregation controls are located below the Meta data information and allow you to aggregate on the time step of your choice using either backward, centred or forward aggregation. Once you have chosen the time step and method you want and clicked the button a pop up box will then appear, letting you choose your aggregation method.



You can either choose the method by individual sensor or if you select the button at the top that will set all sensors to your chosen method. Then press aggregate and your data will be aggregated. If you wish to revert to your original data you can click on the aggregate button again and your data will revert.

Output

The screenshot shows the 'Data Standardizer' application window. The main area contains a data table with columns for various parameters and their values. The right-hand side of the window features a 'Meta information' section with input fields for site details and an 'Output time range' section with date pickers and aggregation options.

Variable	Variable	Variable	Variable	Variable	Variable	Variable	Variable	Variable	Variable
Water Temperature	Water Temperature	Water Temperature	Water Temperature	Dissolved Oxygen	Dissolved Oxygen	Dissolved Oxygen	Dissolved Oxygen	Dissolved Oxygen	Water Temperature
degC	degC	degC	degC	mg/L	mg/L	mg/L	mg/L	mg/L	mg/L
Depth below	Depth below	Depth below	Depth below	Elevation from	Depth below	Depth below	Depth below	Depth below	Depth below
5	7	9	0.2	0.5	10	0.5	10	0.5	0.5
TempWtr_d5(degC)	TempWtr_d7(degC)	TempWtr_d9(degC)	TempWtr_e0.2(degC)	DOsat_d0.5(mg/L)	DOsat_d10(mg/L)	DOconc_d0.5(mg/L)	DOconc_d10(mg/L)	DOconc_d10(mg/L)	TempWtr_d5(degC)

Meta information

Site name: _____
 Site owner: _____
 Contact name: _____
 Phone number: _____
 Contact email: _____
 Organisation: _____
 Elevation (masl): _____
 GPS Grid System: _____
 GPS lat: Decimal GPS long: Decimal
 Country Code: _____
 Site notes: _____
 Data set notes: _____

Output time range

2011-04-27 22:30:00 2011-04-30 23:45:00

Aggregation time step: Quantity
 Direction: Backward Centered Forward

Text to append to filename: .txt

Output file name: (,201104-201104.txt

Include empty lines: Include: Data Metadata file

Export

Outputting in B3 is simple. After you have filled out the meta data fields it will have populated the filename below for you, with the suggested format of Site name(country code)_date range for example Lake Rotoehu(NZ)_201104-201208.txt. Directly below the meta data area is the output range, which lets you output a smaller subset of the data if you wish. You can then pick the format you want below the aggregation box. Then press export and your data will be ready to use in B3 or other software.

Other Features

Below there is a list of the features not previously mentioned, with a quick description on them and how they are used. They are all located on the menu bar under the various categories

Under Data:

Strip incomplete rows

Used to strip any rows from the data that are incomplete or waste rows (over 50% missing)

Header Guess

Tries to auto fill the header boxes based on the details of the current header. Will automatically decode any known headers setup using the Chris McBride method.

Set all data columns to variable

Sets all of the header boxes to the Variable type except for the DateTime column. Note: must be done after date time is standardized into one column.

Under Time:

Standardise date time

Will standardize a date time column into the preferred format of YYYY-MM-DD hh:mm.

Merge and Standardise date time

Will standardize a date time column into the preferred format of YYYY-MM-DD hh:mm as well as merge it from a date and a time column into a singular DateTime column.

Sort Dates

Will organise a dataset by date if it has become scattered and unorganised

Under Options:

Add new parameter

Adds a new parameter to the Chris McBride method for the GLEON control vocabulary. For a list of what is included with data standardizer see Glossary.

Glossary

Standard vocabulary

High-frequency	Abbreviation	Units
Barometric Pressure	BaroP	hPa
Beam Attenuation	BmAtt	1/m
Beam Transmission	BmTran	%
Colored Dissolved Organic Matter	CDOM	ug/L
Chloride	Cl	g/m ³ , mg/m ³
Atmospheric Carbon Dioxide	CO2A	ppm
Conductivity	Cond	mS/cm
Specific Conductance	CondSp	mS/cm
Dissolved Carbon Dioxide	DCO2	mg/L
Dissolved Organic Carbon	DOC	mg/L
Dissolved Oxygen Concentration	DOconc	mg/L
Dissolved Oxygen Saturation	DOsat	%sat
Chlorophyll Fluorescence	FICl	RFU
Phycocyanin Fluorescence	FIPhy	RFU
Heat Flux Evaporative	HFlxEv	W/m ²
Heat Flux Sensible	HFlxSn	W/m ²
Light Attenuation Coefficient	Kd	m ⁻¹
Ammonium	NH4	g/m ³ , mg/m ³
Nitrate	NO3	g/m ³ , mg/m ³
Reduction Oxidation Potential	ORP	mV
pH	pH	
Phosphate	PO4	g/m ³ , mg/m ³
Precipitation Hail	PpHail	hits/cm ²
Precipitation Rainfall	PpRain	mm
Precipitation total	PpT	mm
Longwave Radiation Downwelling	RadLWD	W/m ²
Photosynthetically Active Radiation	RadPAR	umol/m ² /s
Shortwave Radiation Downwelling	RadSWD	W/m ²
Shortwave Light Penetration	RadSWP	
Total Radiation Downwelling	RadTD	W/m ²
Ultraviolet Radiation	RadUV	mW/cm ²
Relative Humidity	RelH	%
Salinity	Sal	-
Precipitation Snow	Snow	mm
Soil Temperature	SoilT	degC
Air Temperature	TmpAir	degC
Dewpoint Temperature	TmpDew	degC
Water Temperature	TmpWtr	degC
Turbidity	Turb	FTU
Datalogger Battery Voltage	VBatLg	V

Radio Battery Voltage	VBatR	V
Vapor Pressure	VP	hPa
Vapor Pressure Deficit	VPDef	hPa
Vapor Pressure Saturation	VPSat	hPa
Solar Panel Voltage	VSol	V
Wave Height	WaveHt	m
Wave Period	WavePd	s
Wind Direction	WndDir	deg
Wind Direction at Peak Speed	WndDrP	deg
Wind Speed Peak	WndMax	deg
Wind Speed Minimum	WndMin	m/s
Wind Speed Average	WndSpd	m/s
Water Column Depth	LvlDpt	m
Water Velocity Horizontal	WtrVIH	m/s
Water Velocity Vertical	WtrVIV	m/s
Flow Volume	FlwVol	m ³
Surface Water Level (Moturiki Datum)	LvlMot	masl