90-e63

Data Management Guidelines for Freshwater Investigations

Prepared by

Robert W. Plotnikoff
John E. Tooley

Washington State Department of Ecology
Environmental Investigations and Laboratory Services Program
Surface Water Investigations Section
Olympia, WA 98504

Data Management

Uniformity in data reporting formats allows access to a much larger base of information. It is the goal of Ecology to incorporate this procedure into the granting process.

The Puget Sound Water Quality Authority (PSWQA) has outlined a format for data storage. This format was designed to include all aspects of field data collection information in addition to water quality laboratory results. The data base was created to address the particular components of environmental monitoring programs in Puget Sound. A long term goal was to enable the incorporation of all data amassed in Puget Sound drainage studies from as many agencies and municipal entities as possible. Information exchange would then be a simple procedure once uniformity in formatted data among these groups become a reality.

Data Format

The PSWQA data format recommendations are described in:

PSWQA. September 1988. Specifications for the Transfer of Data from the Puget Sound Ambient Monitoring Program. Puget Sound Water Quality Authority, Seattle, WA.

A copy of this document may be obtained from:

Puget Sound Water Quality Authority
217 Pine Street Suite 1100
Seattle, WA 98101

Descriptions of each file type and an outline of the file's necessary

components are detailed in this document. It is acceptable to amalgamate files, especially where information is redundant as in SURVEY ID, STATIONID, DATE, TIME, and PERIOD. Water quality studies in drainages of Puget Sound are widely varied and do not all adhere to monitoring the same parameters. With this in mind, file amalgamation is an efficient use of data space.

Files should contain basic information regarding location, date, time, sample collection method, water quality information, and analytical methods. The data entry format outlined by the PSWQA document has been adopted in a Department of Ecology publication, "Guidance for Conducting Water Quality Assessments" prepared in June 1989 (Tables 1 and 2). These tables illustrate the format of a field observations file and a water chemicals/conventionals file. It may be appropriate in some cases to amalgamate information into a single file if the total amount of information is manageable on one spreadsheet.

Once a format has been chosen for a file or files regarding a monitoring program, data entry may commence. The data file will consist of information pertaining to field observations or analysis of water samples collected on a single "survey". A survey refers to the particular sampling session that water samples were continuously collected over the period of one or more days. You should note that the file format includes a column called SURVEY ID. The SURVEY ID is unique to this session within your monitoring program. The remaining column categories are explained in Tables 1 and 2 as well as in the PSWQA (1988) document.

It should be noted that the data file format is left to the discretion of the investigator as long as all required categories are addressed. These required categories are listed in Tables 1 and 2 as

are appropriate definitions of terms. Files can be merged for convenience to form larger data files.

Data Entry Examples

An example of a data spreadsheet with field measurements and laboratory analysis can be found in Table 3. Note that there are two sample dates with thirteen sites and twelve parameters for Anywhere Creek, WA. Not all parameters were measured at each site which explains the reason for blank cells in the spreadsheet. Also, there were no data qualifiers associated with either of the example parameters.

The entry of two parameters, fecal coliform and total phosphorus, into a spreadsheet using the PSWQA format is displayed in Tables 4 and 5. You should first note that the SURVEY ID name may be appropriate for describing the drainage being sampled. The DATE is consistent in a file identified by the unique SURVEY ID. STATION ID and TIME are self explanatory categories. The PERIOD describes the amount of time taken to sample a given station. UPPERDEPTH and LOWERDEPTH describe the vertical range in the water column within which water samples were collected . The VARIABLE refers to the code name for the parameter being sampled. Standardized code names are described for each file category where they are appropriate (Tables 1 and 2). VALUE is the concentration or count of a particular parameter. The QUALIFIER is a code commenting on the numeric data value. In the case of the example data file, there were no qualifiers for the fecal coliform or total phosphorus data reported so the column is blank. SIGNIFICANT DIGITS refers to the number of significant digits in which the value of the parameter was reported. The METHOD is denoted by coded information that is listed in Table 1. Each of the codes describes a method

outlined by a standard guide. QUALITY LEVEL refers to the reliability of the data based on a quality assurance determination. Finally, the UNITS are recorded and are designated for each of the parameters measured.

Each of the data files may contain one or more water quality parameters. The decision on how many parameters are included in each data file is left to the discretion of the investigator. One may wish to separate nutrient data from physical data and place the information in separate data files. One the other hand, it may be more convenient to retain all data in a single file from a particular survey. Each of the data files will eventually be amalgamated into a data base. Creating a data base facilitates different combinations of data queries. Much of the logistics for construction of a central data base is currently being developed by the Puget Sound Water Quality Authority.

Conclusion

This document discussed some of the pertinent topics in creating a data file format following the PSWQA recommendations. Hopefully, this narrative has displayed the flexibility of the data storage system. A large amount of information is presented in the PSWQA (1988) document and does not address data entry possibilities. If there are further enquiries related to data entry format they should be addressed to Surface Water Investigations section of the Washington Department of Ecology.

Table 1. Field Observation File Specifications.

Field Name	Description	Maximum Field Length	Required ?	Codes
SURVEY ID	Identification of monitoring survey	8	Y	
STATION ID	Identifier for station	8	Y	
DATE	Date of observation/sample collection (yymmdd format)	6	Y	
TIME	Time of observation/sample collection (military format)	4	Y	
PERIOD	Period over which sample was collected (hhmm)	4	Y	
UPPER DEPTH	Upper depth where observation was made (nearest .1 m)	6	Y	
LOWER DEPTH	Lower depth where observation was made (nearest .1 m)	6	Y	
TIDE STAGE	Code for tide stage at which observation was made	1	N	1=Ebb 2=Ebb Slack 3=Flood 4=Flood Slack
VARIABLE	Variable measured or observed	10	Y	FLOW=flow DO=dissolved oxygen WTEMPERATUR=temperature, water TURBIDITY=turbidity CONDUCT=conductivity pH=pH WATERDEPTH=water depth
VALUE	Value of variable reported	10	Y	
QUALIFIER	Description to guide in interpretation of data	1	N	
SIGNIFICANT DIGITS	Number of significant digits reported in data value	1	N	
METHOD	Code for methods used	8	Y	OA=Dissolved oxygen-Winkler/Carpenter OB=Dissolved oxygen-Probe/Electrode T1=Turbidity-Turbidometer T2=Turbidity-Transmissometer (1 cm path) T3=Turbidity-Fluorometer T4=Turbidity-Nephelometer T5=Turbidity-Transmissometer (10 cm Path)
QUALITY LEVEL	Quality assurance level assigned to data by reviewer	1	N	1=Data collected in accordance with Puget Sound Protocols or methods acceptable for PSAMP and there are no data quality problems 2=Same as above except problems arose and were corrected

Table 1. (Continued)

Field Name	Description	Maximum Field Length	Required ?	Codes
QUALITY LEVEL (con't)				3=Data was not collected in accordance with protocols or quality control problems could not be corrected 4=Data was lost
UNITS	Units in which data value is reported	2	Y	MS=ppm (mg/kg) ML=ppm (mg/L) DC=degrees celsius UC=umhos/cm MC=meters/sec PH=pH
MEASUREMENT BASIS	Weight basis for data value measurement (wet or dry)	1	N	D=dry weight W=wet weight

Table 2. Water Chemicals/Conventionals File Specifications.

Field Name	Description	Maximum Field Length	Required ?	Codes
SURVEY ID	Identification of monitoring survey	8	Y	
STATION ID	Identifier for station	8	Y	
DATE	Date of observation/sample collection (yymmdd format)	6	Y	
TIME	Time of observation/sample collection (military format)	4	Y	
PERIOD	Period over which sample was collected (hhmm)	4	Y	
UPPER DEPTH	Upper depth where observation was made (nearest .1 m)	6	Y	
LOWER DEPTH	Lower depth where observation was made (nearest .1 m)	6	Y	
TIDE STAGE	Code for tide stage at which observation was made	1	N	1=Ebb 2=Ebb Slack 3=Flood 4=Flood Slack
VARIABLE	Variable measured or observed	10	Y	AMMONIA=Ammonia, Total (ug/L) PHOSPHATE=Phosphorus, Total (ug/L) ORTHO PHOS=Ortho Phosphorus (ug/L) NO3-N=Nitrate (ug/L) NO2-N=Nitrite (ug/L) TOT SOLIDS=Total Suspended Solids (mg/L) ALUMINUM=Aluminum (ug/L) ANTIMONY=Antimony (ug/L) ARSENIC=Arsenic (ug/L) CADMIUM=Cadmium (ug/L) CHROMIUM=Chromium (ug/L) COPPER=Copper, Total (ug/L) IRON=Iron (ug/L) LEAD=Lead (ug/L) MANGANESE=Manganese (ug/L) MERCURY=Mercury, Total (ug/L) NICKEL=Nickel, Total (ug/L) SILVER=Silver (ug/L) ZINC=Zinc (ug/L) CATIONS=Cations (mg/L) ANIONS=Anions (ng/L) FECALCOLIF=Fecal Coliform (MPN/1000 mL) HARDNESS=Total Hardness (mg/L) ALKLNTY=Alkalinity (mg/L)

Table 2. (Continued)

Field Name	Description	Maximum Field Length	Required ?	Codes
VALUE	Value of variable reported	10	Y	
QUALIFIER	Description to guide in interpretation of data	1	N	
SIGNIFICANT DIGITS	Number of significant digits reported in data value	1	N	
METHOD	Code for methods used	8	Y	P8603CS=Recommended methods for analysis of sediment conventionals SM85CW=Standard Methods (APHA 1985) P8608M-CVAA=Cold vapor atomic absorption spectrometry P8608M-GFAA=Graphite furnace atomic absorption spectrometry P8608M-ICP=Inductively coupled plasma emmission spectroscopy P8608M-HGAA=Hydride generation atomic absorption P8610F-SW=Recommended methods for fecal coliform analysis in water or sediment
QUALITY LEVEL	Quality assurance level assigned to data by reviewer	1	N	1=Data collected in accordance with Puget Sound Protocols or methods acceptable for PSAMP and there are no data quality problems 2=Same as above except problems arose and were corrected 3=Data was not collected in accordance with protocols or quality control problems could not be corrected 4=Data was lost
UNITS	Units in which data value is reported	2	Y	MS=ppm (mg/kg) ML=ppm (mg/L) DC=degrees celsius UC=umhos/cm MC=meters/sec PH=pH
MEASUREMENT BASIS	Weight basis for data value measurement (wet or dry)	1	N	D=dry weight W=wet weight

Table 3. Raw data table for Anywhere Creek, WA.

Anywhere Creek, WA

Date	Station	Temp (oC)	Cond (umhos)	pH	DO (mg/L)	Flow (CFS)	Turb (NTU)	TSS (mg/L)	NO3+NO2 (mg N/L)	NH3 (mg N/L)	TP (mg P/L)	FC (#/100 mL)	BOD5 (mg O/L)	
12/21/88	P1.3	5.25	171	6.99	8.7	46.67	3	6	2.600	0.530	0.200	20	3	U
12/21/88	UN2.0	5.69	105	6.60	9.0	3.89	2	4	0.970	0.050	0.130	520		
12/21/88	P2.5	5.51	186	6.99	5.3			2				220		
12/21/88	P3.7	5.62	172	6.94	6.0	36.75	2	6	2.200	0.370	0.340	530	3	
12/21/88	P5.2	5.47	171	6.97	8.5	20.18	3	4	2.000	1.200	0.370	250		
12/21/88	UN5.4	5.70	201	7.06	6.1			2				32		
12/21/88	P6.0	6.73	217	7.29	8.4			3				66		
12/21/88	P7.1	4.77	81	7.23	12.7	11.15	2	4	2.000	0.020	0.120	49	3	U
12/21/88	P7.3	4.63	82	7.15	12.4	5.30	3	2	1.800	0.020	0.100	66		
12/21/88	UN8.0	6.76	154	7.01	10.8			4				3		
12/21/88	F0.8	4.32	109	7.18	12.3	14.69	3	8	2.500	0.030	0.150	170	3	U
12/21/88	F2.5	4.58	106	7.00	11.9			9				170		
12/21/88	F4.3	5.52	102	7.19	11.7			8				20		
01/09/89	P1.3	4.44	140	6.77	9.4	72.95	3	26	2.900	0.080	0.170	75	3	K
01/09/89	UN2.0	5.00	100	6.24	9.5	4.85	2	7	1.200	0.050	0.150	26		
01/09/89	P2.5	4.82	167	6.75	6.3			2				110		
01/09/89	P3.7	4.71	155	6.72	7.2	34.71	2	7	2.200	0.200	0.320	365	3	K
01/09/89	P5.2	4.86	157	6.81	9.1	29.15	3	6	2.200	1.000	0.440	240		
01/09/89	UN5.4	5.33	207	6.92	6.0			2				200		
01/09/89	P6.0	6.46	210	7.20	8.8			3				80		
01/09/89	P7.1	4.21	76	7.09	12.9	18.65	3	6	2.100	0.010	0.120	210	3	K
01/09/89	P7.3	4.11	75	6.68	12.4	9.74	3	5	2.100	0.110	0.220	210		
01/09/89	P8.1	4.11	63	7.29	12.8			5				63		
01/09/89	F0.8	3.54	97	6.84	12.4	29.45	3	28	2.500	0.060	0.150	330	3	K
01/09/89	F2.5	4.21	99	7.00	12.0			12				120		
01/09/89	F4.3	4.80	95	7.20	12.2			9				6		

Table 4. Formatted data file following the Puget Sound Ambient Monitoring Program recommendations for standardization of data reporting.

File no. 1

SURVEYID	STATIONID	DATE	TIME	PERIOD	UPPERDEPTH	LOWERDEPTH	VARIABLE	VALUE	QUALIFIER	SIGNIF	METHOD	QALEVEL	UNITS	VARIABLE	VALUE	QUALIFIER	SIGNIF	METHOD	QALEVEL	UNITS
AnyCrk1	P1.3	881221	0833	0015	0.1	0.3	FECALCOLIF	20		2	SM85CW	1	MPN/100ml	PHOSPHATE	0.200		3	P8603CS	1	ML
AnyCrk1	UN2.0	881221	1053	0010	0.1	0.3	FECALCOLIF	520		2	SM85CW	1	MPN/100ml	PHOSPHATE	0.130		3	P8603CS	1	ML
AnyCrk1	P2.5	881221	1130	0010	0.1	0.3	FECALCOLIF	220		2	SM85CW	1	MPN/100ml							
AnyCrk1	P3.7	881221	0924	0015	0.1	0.3	FECALCOLIF	530		2	SM85CW	1	MPN/100ml	PHOSPHATE	0.340		3	P8603CS	1	ML
AnyCrk1	P5.2	881221	1310	0015	0.1	0.3	FECALCOLIF	250		2	SM85CW	1	MPN/100ml	PHOSPHATE	0.370		3	P8603CS	1	ML
AnyCrk1	UN5.4	881221	1325	0010	0.1	0.3	FECALCOLIF	32		2	SM85CW	1	MPN/100ml							
AnyCrk1	P6.0	881221	1345	0015	0.1	0.3	FECALCOLIF	66		2	SM85CW	1	MPN/100ml							
AnyCrk1	P7.1	881221	1500	0015	0.1	0.3	FECALCOLIF	49		2	SM85CW	1	MPN/100ml	PHOSPHATE	0.120		3	P8603CS	1	ML
AnyCrk1	P7.3	881221	1415	0015	0.1	0.3	FECALCOLIF	66		2	SM85CW	1	MPN/100ml	PHOSPHATE	0.100		3	P8603CS	1	ML
AnyCrk1	UN8.0	881221	1545	0010	0.1	0.3	FECALCOLIF	3		2	SM85CW	1	MPN/100ml							
AnyCrk1	F0.8	881221	1010	0015	0.1	0.3	FECALCOLIF	170		2	SM85CW	1	MPN/100ml	PHOSPHATE	0.150		3	P8603CS	1	ML

Table 5. Formatted data file following the Puget Sound Ambient Monitoring Program recommendations for standardization of data reporting.

File no. 2

SURVEYID	STATIONID	DATE	TIME	PERIOD	UPPERDEPTH	LOWERDEPTH	VARIABLE	VALUE	QUALIFIER	SIGNIF	METHOD	QALEVEL	UNITS	VARIABLE	VALUE	QUALIFIER	SIGNIF	METHOD	QALEVEL	UNITS
AnyCrk2	P1.3	890109	0855	0015	0.1	0.3	FECALCOLIF	75		2	SM85CW	1	MPN/100ml	PHOSPHATE	0.170		3	P8603CS	1	ML
AnyCrk2	UN2.0	890109	1230	0010	0.1	0.3	FECALCOLIF	26		2	SM85CW	1	MPN/100ml	PHOSPHATE	0.150		3	P8603CS	1	ML
AnyCrk2	P2.5	890109	1125	0010	0.1	0.3	FECALCOLIF	110		2	SM85CW	1	MPN/100ml							
AnyCrk2	P3.7	890109	1030	0015	0.1	0.3	FECALCOLIF	370		2	SM85CW	1	MPN/100ml	PHOSPHATE	0.320		3	P8603CS	1	ML
AnyCrk2	P5.2	890109	1430	0015	0.1	0.3	FECALCOLIF	240		2	SM85CW	1	MPN/100ml	PHOSPHATE	0.440		3	P8603CS	1	ML
AnyCrk2	UN5.4	890109	1420	0010	0.1	0.3	FECALCOLIF	200		2	SM85CW	1	MPN/100ml							
AnyCrk2	P6.0	890109	1450	0015	0.1	0.3	FECALCOLIF	80		2	SM85CW	1	MPN/100ml							
AnyCrk2	P7.1	890109	1510	0015	0.1	0.3	FECALCOLIF	210		2	SM85CW	1	MPN/100ml	PHOSPHATE	0.120		3	P8603CS	1	ML
AnyCrk2	P7.3	890109	1600	0015	0.1	0.3	FECALCOLIF	210		2	SM85CW	1	MPN/100ml	PHOSPHATE	0.220		3	P8603CS	1	ML
AnyCrk2	UN8.0	890109	1552	0010	0.1	0.3	FECALCOLIF	63		2	SM85CW	1	MPN/100ml							
AnyCrk2	F0.8	890109	1147	0015	0.1	0.3	FECALCOLIF	330		2	SM85CW	1	MPN/100ml	PHOSPHATE	0.150		3	P8603CS	1	ML