Data Dictionary

How to understand each part of the StreamWatch data

Data Condition

Data conditions refer to the quality assurance status of data, indicating the extent to which the data has been reviewed or verified. Some datasets may be professionally checked and validated to varying degrees, while others remain unverified. Although unverified data can still be valid, it is essential to note its review status before drawing scientific conclusions.

Accepted	Data has been reviewed and meets quality standards without requiring changes.		
Corrected	Data contained issues that have been identified and fixed.		
Corrected; Erroneous	Data was found to be incorrect, corrected, but still flagged due to serious past errors.	**Corrected; Incomplete**	Incomplete data that has been partially fixed or supplemented.
Duplicate	Repeated data entries that do not provide new information.	**Erroneous**	Data contains known errors and is unreliable without correction.
Flagged	Data is marked for review due to potential issues or inconsistencies.	**Incomplete**	Data is missing key elements or measurements.
Minor Deviation	Data shows small inconsistencies but is generally usable.	**Minor Deviation; Corrected**	Slight issues were found and corrected.
Moderate Deviation	Data shows noticeable inconsistencies that may affect reliability.	**Provisional**	Data is preliminary and subject to change upon further review.
Site Conflict	Data conflicts with other data from the same site or location.	**Unchecked**	Data has not yet been reviewed or verified for accuracy.

Notes

This is where important information unique to each entry can be entered if there is no established location on the spreadsheet. This may also include information regarding the data condition, such as clarifying what corrections have been started, completed, or need to be done.

Method

The method describes how the data presented was collected, which can provide helpful information on certain aspects, such as the accuracy, of the entry.

BACT	These samples are collected by streamwatch volunteers, and then processed to monitor E.Coli, Turbidity, and Nutrients in order to assess water quality and safety standards.
BAT	StreamWatch BATS monitors the populations of organisms called macroinvertebrates. Water quality ratings are based on the abundance of various species and their sensitivity to pollutants.
Hanna	This data was collected using a Hannameter, which is a digital tool used to assess water quality.
LaMotte	This data was collected using a LaMotte kit, in which volunteers asses quality of water manually.

Streamwatch Site Locations

Code	Lat	Lon
AC1	40.21823	-74.7631
AC1a	40.21816	-74.7645
AC2	40.23092	-74.7422
AC3	40.23455	-74.7388
AC4	40.264	-74.7114
AC5	40.2699	-74.6718
AC6	40.24192	-74.5948
AC7	40.21846	-74.5602
AC8	40.20721	-74.5499
AC9	40.24577	-74.7302
APL1	40.21546	-74.5175
ATL1	40.17546	-74.5866
AWL1	40.40017	-74.843
AX1	40.37861	-74.9489
AX2	40.40322	-74.9139
BB1	40.31836	-74.6125
BB2	40.27821	-74.5771
BD1	40.4198	-74.6392
BD1a	40.41603	-74.643
BD2	40.41563	-74.6638
BD3	40.39454	-74.7009
BD4	40.38429	-74.7406
BD5	40.38612	-74.7607
BD6	40.42158	-74.6324
BK1	40.43601	-74.6563
BLDL1	40.34163	-74.784
BRL1	40.31083	-74.5121
CAN1	40.43613	-74.6146
CAN-DP	40.33164	-74.6526
CAN-LM	40.52874	-74.582
CB1	40.32665	-74.6029
CB2	40.30979	-74.5168
CB3	40.32612	-74.6035
CB4	40.30514	-74.4732
CB5	40.32993	-74.5456
CC1	40.13452	-74.5584
CH1	40.38274	-74.6848
CL1	40.37394	-74.62
CL2	40.34648	-74.6274
CL4	40.34174	-74.6475
CL5	40.36686	-74.6257
COL1	40.25536	-74.7239
CR1	40.45709	-74.6856
CR2	40.45434	-74.6604
CRL1	40.32272	-74.7763
CTL1	40.28061	-74.7397
CVL1	40.27285	-74.7795
DAM-6	40.53026	-74.5883
DAM-7	40.53107	-74.5872
DB1	40.33388	-74.6117
DB2	40.33122	-74.6165
DBA	40.36185	-74.5451
DBB	40.35129	-74.5701
DE1	40.36052	-74.9454
DP3	40.30679	-74.6685
DS21-1	40.24411	-74.656
EB1	40.21675	-74.4792
Ely		
ETL1	40.25342	-74.4991
FC1	40.31666	-74.882
GMP1	40.31296	-74.6063
GRPL1	40.18783	-74.6886
HA1	40.36599	-74.6275
HA2	40.35905	-74.6387

Site

The location of the waterbody each sample was taken from is represented as a code. In general, each letter combination is a unique water body (**AC**1) . Additional numbers following a code (AC**1**) distinguish specific areas within the water body. Refer to StreamWatch site Locations table.

Date / Time

The date the sample was collected is listen in Day of the Week, Month, Day, Year format. The time the sample was taken is listed in 2400 format.

Air Temperature

Temperature of the air at the collected site measured in celsius.

Water Temperature

Temperature of the water at the collected site measured in celsius.

Nitrate

This column showcases the concentration of nitrate in the samples , measured in micrograms per liter.

Nitrate is a chemical compound with the formula NO_3^-.

It's commonly found in fertilizers, agricultural runoff, sewage, and natural soils.

NIT4_GL

For data collected using the LaMotte method, a note is made indicating if the results of the data is above or below the kits detection limit. A detection limit is the lowest concentration of a substance that can be reliably distinguished from zero.

<	Indicates that this measurement was below the detection limit.
=	Indicates that this measurement was within the detection limit.
>	Indicates that this measurement was greater than the detection limit.

Phosphates

This column showcases the concentration of phosphates in the samples, measured in milligrams per liter.

Phosphate is a chemical containing phosphorus

It occurs naturally but is also added to water and soil through fertilizers, detergents, and sewage. In water, excessive phosphate can cause eutrophication—an overgrowth of algae that depletes oxygen and harms aquatic life.

PHOS_GL

Refer to the NIT4_GL section.

HB1		
HB2		
HB3	40.37076	-74.6192
HL1	40.34799	-74.7472
HL2	40.34624	-74.7443
HL3		
HO1	40.34053	-74.7431
HO2	40.35743	-74.7656
HO3	40.36119	-74.767
HO3a	40.36028	-74.7679
JC0	40.34156	-74.8385
JC1	40.33163	-74.8361
JC2	40.32816	-74.8364
JC3	40.31851	-74.838
JC4	40.31363	-74.8375
JC4A		
JC4T		
JC5	40.30622	-74.8342
JC6	40.29863	-74.8349
JC9	40.2802	-74.8509
LSC1	40.27023	-74.7153
LSP1	40.29327	-74.7254
MB1	40.34885	-74.6878
MB2	40.3651	-74.6768
MB3	40.36373	-74.6802
MB5	40.36775	-74.6643
MDWL1	40.25812	-74.5202
MI2	40.24396	-74.6595
MI3	40.24918	-74.711
MI4	40.24687	-74.6859
MI4b	40.24686	-74.6876
MO1	40.32442	-74.9169
MR1	40.33539	-74.6241
MR2	40.32199	-74.6079
MR3	40.439	-74.6178
MR4	40.40685	-74.6365
MR5	40.47458	-74.5754
MR6	40.47537	-74.5763
MR7	40.50369	-74.5873
MR8	40.54222	-74.568
MRA	40.20887	-74.3939
MRB	40.26158	-74.42
MRC	40.27453	-74.4721
MRCL1	40.26578	-74.6417
MRD	40.29312	-74.5269
PB1	40.3924	-74.8351
PL1	40.2702	-74.5221
PMRL1	40.36323	-74.6768
PNVL1	40.22672	-74.438
PP1	40.32762	-74.6009
PR1	40.44366	-74.6466
PR2	40.41981	-74.6406
PR2a	40.4265	-74.6437
PR3	40.46807	-74.6488
PRP1	40.34999	-74.5609
PRP1a	40.34977	-74.5646
RB1	40.23207	-74.4617
RB2	40.28306	-74.5369
RBA	40.27351	-74.5259
RK1	40.4197	-74.7219
RK2	40.41299	-74.6797
RO1	40.50206	-74.6257
RO1a	40.50546	-74.6236
RO2	40.5322	-74.5895

pH

pH is a measure of how acidic or basic (alkaline) a substance is, on a scale from 0 to 14. It tells us how many hydrogen ions (H^+) are present in a solution. pH matters because it affects solubility and availability of nutrients and metals which influences aquatic organisms' health—most can only survive in a narrow pH range. Shifts in pH can indicate pollution, chemical discharge, or biological activity.

PH_GL

Refer to the NIT4_GL section.

Turbidity

Turbidity is a measure of how clear or cloudy water is, caused by the presence of suspended particles such as silt, clay, organic matter, algae, and microscopic organisms that scatter and absorb light. High turbidity can reduce light penetration, affecting photosynthesis of aquatic plants, raise water temperature by absorbing more sunlight, clog fish gills and harm aquatic organisms. High turbidity often indicates erosion, runoff, algal blooms, or pollution.

TURB_GL

Refer to the NIT4_GL section.

DO ppm / %DO

Dissolved oxygen is the amount of gaseous oxygen (O_2) present in water, available for aquatic organisms to use for respiration. Fish, macroinvertebrates, and aerobic bacteria require DO to survive. DO is measured two ways, in parts per milion (1mg/Liter), or in percent saturation.

DO ppm

DO ppm stands for Dissolved Oxygen in parts per million.
It measures the absolute concentration of oxygen dissolved in water.
Measured in mg/L

%DO

%DO is the percentage of oxygen saturation relative to the maximum amount that water can hold at a given temperature, salinity, and pressure.
For example, 100% DO means water is holding all the oxygen it can at that temperature.

DO is a direct indicator of water quality. High DO generally means good conditions for aquatic life, while low DO suggests pollution, high biological oxygen demand, or poor aeration. Low DO can indicate excessive organic pollution oralgal blooms, where DO rises during the day (photosynthesis) and drops drastically at night (respiration).

Conductivity

Conductivity is a measure of water's ability to conduct electrical current. It depends on the concentration of dissolved ions (e.g. sodium, chloride, nitrate, sulfate, calcium, magnesium). Measured in µS/cm . Sudden increase in conductivity can signal: sewage or industrial discharge rich in ions, agricultural runoff containing fertilizers (nitrate, phosphate salts)

Chloride

Chloride (Cl^-) is a negatively charged ion (anion) found commonly in natural waters. We measure chloride in mg/L
It originates from dissolved salts, mainly sodium chloride (NaCl) but also potassium, calcium, or magnesium chloride. Chloride can originate from natural sources, or be causes by human action through road de-icing salts, sewage and wastewater discharges, industrial effluents, and gricultural runoff. Chloride is an indicator of pollution.

N. Bat

This number represents the number of bugs found by volunteers during a particular day at the specified site. This is used to assess stream health.

ROL1	40.33083	-74.7575
SB1	40.333	-74.682
SB2	40.35282	-74.7104
SB3	40.35855	-74.6911
SB4	40.33175	-74.7783
SB4b	40.33328	-74.782
SB5	40.37422	-74.7939
SB6	40.33136	-74.767
SB9	40.33097	-74.6542
SBA	40.34959	-74.7826
SBB	40.40146	-74.8046
SBC	40.38804	-74.8273
SBD	40.39249	-74.8408
SBE	40.40341	-74.8547
SC1	40.24883	-74.7562
SC2	40.25722	-74.7296
SC3	40.28491	-74.7664
SC4	40.26318	-74.7596
SC5	40.25424	-74.7217
SC6	40.25696	-74.7497
SHA	40.34677	-74.5573
SM1	40.47282	-74.5709
SM2	40.46973	-74.5442
SM3	40.45537	-74.5145
SPBP1	40.50947	-74.5654
SPBP2	40.51015	-74.5606
SW1	40.36404	-74.945
SYL1	40.27152	-74.7753
SYNL1	40.41237	-74.6981
TM1	40.45651	-74.5857
WB1	40.3088	-74.8829
WB2A	40.30714	-74.8338
WB2B		
WB2T		
WP1	40.36051	-74.7667
WP2	40.35844	-74.7661
WP3	40.35905	-74.7656
WP4	40.35805	-74.7659

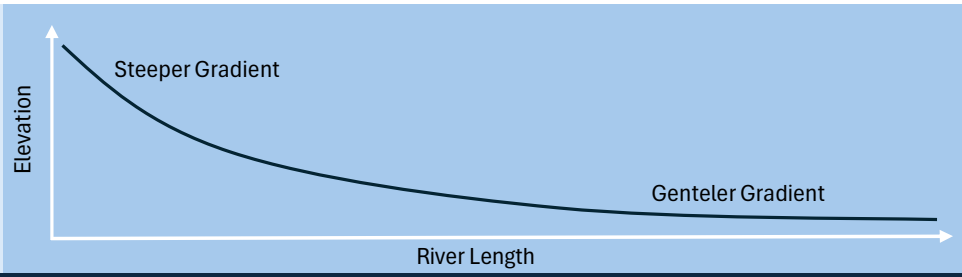
NJIS

New Jersey impairment score, measurment of water quality based upon bug counts. Lower scores represent higher habitat impairment.

HGMI.Family

This value represents the number of Ephemeroptera, Plecoptera, and Trichoptera, insect taxonomic orders for mayflies, stoneflies, and caddisflies. The relative abudence of these insect families can provide information about habitat health. HGMI stands for high gradient macroinvertibrate index.

What is a high gradient stream? A high gradient stream is characterized by a steep slope that causes rapid water flow, resulting in a high chance of eroison.



HGMI.Genus

This column is very similar to the HGMI.Family column, but instead calculates a score based on the recorded genus instead of family. Utilizing genus instead of species allows a more in depth data, as it is a more specific measurment.

CPMI

The CPMI score is the costal plains macroinvertibrate index, which provides similar information to the HGMI, but in low gradient environments .

FBI

FBI stands for family biotic index, which is a subsection of the HGMI.family measurment. The FBI assigns a tolerance value to each family of macroinvertebrates, with higher values indicating greater sensitivity to pollution.

E. Coli Mod.

Indecaates how the E. Coli score fits into the limit of decetion. Refer to NIT4_GL section for further explaination of the <, =, > symbols.

E. Coli Result

Escherichia coli (E. coli) is a type of bacteria found in the intestines of humans and warm-blooded animals. Most strains are harmless, but some can cause illness. E. coli gets into streams through fecal contamination such as runoff from farms with livestock manure, leaking septic tanks or sewage systems and stormwater runoff carrying pet waste. The presence of E. coli in streams is used as an indicator of faecal pollution. It does not necessarily mean harmful bacteria are present, but it signals that pathogens from faeces may be in the water. E. Coli levels are tested to assess water safety for recreation (e.g. swimming, kayaking), determine drinking water risk if the stream feeds supplies, and to monitor ecosystem health. High E. coli levels suggest the possible presence of pathogens that can cause gastrointestinal illness, ear, eye, or skin infections, and more severe infections from certain strains (e.g. E. coli O157:H7). Safe levles are usually <235 CFU/100 mL (in the US EPA guidelines). We measure ecoli in MPN--most probable colonies per 100 mL.

Epifanual Substrate/ Available Cover

Epifaunal substrate refers to the hard surfaces within an aquatic environment that serve as a habitat for epifauna, which are organisms that live on or attached to the bottom of a water body.

Includes:	rocks, sand, gravel, logs, and other solid surfaces.

Embeddedness - High Gradient

Embeddedness is the degree to which larger substrate particles (like gravel, cobble, and boulders) are surrounded or buried by finer sediments (sand and silt)

Pool Substrate Characterization - Low Gradient

This metric focuses on the types and condition of the bottom materials within pools and their impact on aquatic life. Low gradient streams typically have substrates dominated by finer sediments like sand, mud, or clay, with less coarse material like gravel or cobble.

Velocity/ Depth Combination - High Gradient

In high-gradient or steep channels, the combination of water velocity and depth significantly impacts habitat diversity and overall stream health. High gradient channels are characterized by relatively steep slopes, often leading to supercritical flow, where the water velocity exceeds the speed of a small wave on the surface,

Pool Variability - Low Gradient

In low-gradient streams, pool variability refers to the variety of pool types present, including large-shallow, large-deep, small-shallow, and small-deep pools. A diverse mix of pool types is important for supporting a wide range of aquatic species. Low-gradient streams with low sinuosity and monotonous pool characteristics may not offer sufficient habitat diversity to sustain a healthy aquatic community.

Sediment Deposition

Sediment deposition is the process by which particles such as sand, silt, clay, gravel, and organic material carried by the stream are laid down or settle out of the water column and accumulate on the streambed or floodplain. Larger numbers represent optimal conditions (little/no new sediment deposits), whilst smaller numbers represent poor conditions (large amounts of new sediment deposits).

Channel Flow Status

Channel flow status refers to the amount and continuity of water flow within the stream channel at a given time. It describes how much of the streambed is covered by flowing water and whether the flow is continuous or intermittent. High numbers represent high continuous water flow, whilst low numbers represent little/no flow and standing pools.

Channel Alteration

Channel alteration refers to any change to the natural shape, structure, path, or flow characteristics of a stream channel. These changes can be natural (e.g. erosion, sediment deposition) or caused by human activities (e.g. straightening, dredging). High scores represent normal stream patterns, while low scores represent streams shored by structures/cement and over 80% channelized.

Frequency of Riffles - High Gradient

Riffles are shallow sections of a stream or river where water flows swiftly over a coarse substrate of gravel, cobble, or boulders, creating turbulence and surface disturbance (rippling or “white water”). High scores represent continuous and frequency riffles, and low scores repreent streams wirh mostly fla or shallow water and litle to no riffles.

Channel Sinuosity - Low Gradient

Channel sinuosity is a measure of how much a stream meanders or curves across its floodplain. It is calculated as

Sinuosity = Channel length / Valley length

Streams who become 3-4x longer with curves than if they were straight receive high scores, while those who do not notbaly increase with curves recieve lower scores.

Bank Stability - Left / Right

Bank stability refers to the ability of a streambank to resist erosion and maintain its shape, structure, and position over time under the forces of flowing water, gravity, and other environmental factors. High numbers represent more stable banks.

Factors effecting bank stability:

Vegetation Cover	Roots bind soil, increasing cohesion and resistance to erosion. Vegetation absorbs water, reducing soil saturation and slumping risk. Lack of vegetation reduces stability.
Soil Type and Composition	Clay and cohesive soils resist erosion better than sandy, non-cohesive soils. Mixed substrates with gravels and cobbles provide armoring against flow.
Bank Slope	Steeper banks (>45°) are more prone to collapse, especially when undercut. Gentle-sloped banks are more stable.
Hydrology and Flow Regime	High flow events or rapid fluctuations (e.g. from dams) increase erosion risk. Baseflow supports vegetation growth, improving stability.
Human Activities	Removal of riparian vegetation, livestock trampling, construction, or channel alterations destabilize banks. Bank armoring can prevent erosion but may shift problems downstream.
Groundwater Influence	Saturated banks are heavier and weaker, increasing slumping or mass wasting risk. Fluctuating groundwater levels can undermine bank material cohesion.

Bank Vegetative Protection - Left / Right

Bank vegetative protection refers to the use of plants and vegetation cover to stabilize streambanks, reduce erosion, and maintain the integrity of the channel edge. Higher scores indicate streams with over 90% of the bank covered by native vegetation. Example include:

Grasses and Herbaceous Plants: Fast-growing with fibrous roots effective for surface erosion control.

Shrubs: Deeper root systems than grasses, enhancing bank stability.

Trees: Deep, strong roots provide significant structural reinforcement.

Combined Riparian Vegetation: The most effective protection comes from multi-layered vegetation (grasses + shrubs + trees), offering surface cover, flow resistance, and deep root stabilization.

Riparian Vegetative Zone – Left / Right

The riparian vegetative zone is the area of vegetation directly adjacent to a stream, river, lake, or other waterbody. It includes trees, shrubs, grasses, and other plants growing along the banks and floodplain. This zone forms the transitional area between aquatic and upland ecosystems. This zone is important in Bank stabilization, water quality protection, habitat provision, temperature regulation, flood mitigation, and nutrient cycling. High score represent streams with large riparian vegetative zones, with human activities having little ipact on the zone.