

IOOS Conventions for TSV Encoding

Editor: Jeff de La Beaujardière, NOAA/NOS/IOOS
Contributors: Mike Garcia, NOAA/NWS/NDBC

1 Introduction

This document describes the conventions used by the Integrated Ocean Observing System (IOOS) program to encode observation data as plain text with tab-separated values (TSV). TSV means that the data are expressed as a sequence of multiple data values or metadata attributes separated by tabs (ASCII character 0x09) on a single line of text. Each line of text represents a single point in time and space. Multiple lines are used for additional times or locations. TSV data can be thought of as an array of rows (one per line) and columns (one per value in each line).

This document applies to the following data value types: Scalar, Vector, Multivalued, and Spectral. This document applies to the following Sampling Feature types: Point, Vertical Profile, Horizontal Profile, 2D Trajectory, 3D Trajectory, and Collection.

NOTE: This document does not address conventions for reporting multiple phenomena from different sensors from a single station in a single response. This remains an area to be discussed.

NOTE: This document does not discuss other sampling feature types including regular grids, irregular grids, unstructured grids, or volumetric data. IOOS does not encode such data as TSV—instead, the binary NetCDF format with CF conventions is used.

These conventions are intended to be applied by the IOOS Sensor Observation Service (SOS) instances, but could be used to transmit and store TSV-encoded data from other sources as well.

1.1 Revision History

1.0.0: Original version. Published 2010-06-09.

1.0.0b: Minor editing; no change to specification. 2010-07-20.

1.0.0c: Update sample URLs to use production server instead of test server. 2010-12-20.

2 General TSV Conventions

2.1 Basic Structure

IOOS TSV responses shall use the following basic structure:

- a line break (CR/LF) between each line;

- a TAB between each value;
- no TABs permitted in any value;
- two TABs in a row to designate a missing value;
- no TAB after the last value, unless the last value is missing in which case the preceding TAB is retained;
- all characters within a value are significant (including spaces, commas, and quotes).

2.2 MIME Type

IOOS TSV files shall use the MIME type registered with the Internet Assigned Numbers Authority (IANA, 2010):

```
text/tab-separated-values
```

The MIME type shall be indicated by the originating server using the HTTP Content-Type header.

NOTE: The NDBC SOS server is presently using the HTTP Content-Disposition header to specify a filename and open a dialog box. We are considering conventions to make the resulting filename more reflective of the phenomenon, procedure and time of the data inside.

2.3 Compression

IOOS servers shall offer uncompressed TSV. Servers may also offer compressed TSV as well (using gzip, ZIP, etc.) but this is not required. If the TSV result is compressed the originating server shall indicate this fact using the HTTP Content-Encoding header. Clients should use the HTTP Accept-Encoding mechanism to request compression if desired, being prepared to handle either compressed or uncompressed responses.

2.4 Number and order of values

Every line in a given TSV response shall have the same number and order of values, with missing values indicated by an empty field (two tabs) or by agreed-upon terms to indicate missing information (e.g., 'missing' or 'N/A' or 'NULL') in some cases.

2.5 Header row and initial columns

The first line of every IOOS TSV response shall comprise a list of column headers that provide names for the values in the following rows. If a value has a unit, it shall be specified in square brackets [] at the end of the column name (no text is permitted after the closing bracket). This line is known as the "header row." All other lines are referred to as "data rows."

The first several fields of the header row shall be, in this order:

```
station_id:METAVAR:TEXT:61    sensor_id:METAVAR:TEXT:61    latitude [degree]
      longitude [degree]    time_ISO8601    depth [m]
```

The initial fields of each data row shall provide values for these quantities.

`station_id:METAVAR:TEXT:61` and `sensor_id:METAVAR:TEXT:61` represent the `station_id` and `sensor_id`, URNs as defined in the IOOS Convention for URN Identifiers. These fields are named to provide consistency with ODV tools.

Latitude and longitude are in decimal degrees relative to World Geodetic System 1984 (WGS84) coordinate reference system (CRS ID urn:ogc:def:crs:epsg::4326).

`time_ISO8601` is in ISO 8601 format with punctuation (normally yyyy-mm-ddThh:mm:ssZ, with variations for non-specific times (e.g., climatological average of temperature in August regardless of year).

Depth is in meters, positive below the surface of the water.

NOTE: The foregoing implies that all data values in a given row are from the same station, sensor, location and time.

2.6 Phenomenon-specific columns

As discussed in Section 3 below, the remaining fields will depend on the phenomenon, data value type, and sampling feature types. In order to ensure compatibility of TSV encodings from different servers, the ordering of some mandatory fields for each will be specified in this document for various phenomena. Data providers may offer additional fields after the mandatory fields; their order is not specified here but a name shall be provided for every column in the TSV response.

2.7 Use of CF Names

In the header row, names of data values shall use the Climate and Forecast (CF) Standard Names where possible. The units of the quantity are not required to be the same as the "canonical" CF unit as long as there is a well-known conversion formula from the specified units to the canonical units. Example: The canonical unit of Temperature in CF is Kelvin, but data may be reported in Celsius.

2.8 Time-dependent metadata

Some types of data may include time-dependent metadata associated with each measurement. Example: data on ocean currents that includes pitch, roll and yaw information for the sensor. The general practice shall be to place the principal data values first in each data row, followed by any associated time-dependent metadata.

2.9 Result Size

IOOS does not impose limits on the number of lines in the TSV result or values on each line. However, we note that common spreadsheet applications have a limit of 256 values per row and 65536 rows per sheet.

The TSV result may be compressed (see above under General conventions) to minimize transmitted file size.

2.10 Empty Dataset

A TSV response to a valid request that yielded no data at all (e.g., time range or bounding box did not match any stations) shall contain a header row and zero data rows.

NOTE: This approach differs from the handling of empty datasets in the IOOS GML case. There, if no data corresponds to a query, an OGC Web Service Exception is returned (i.e., a brief XML document explaining the nature of the problem). The Empty Dataset response proposed here means these TSV conventions could be applied regardless of whether the data are served by SOS or obtained through other means.

2.11 Sort Order

Data from multiple times from a single station shall be sorted in order from earliest time to latest time.

Data from multiple depths from a single station shall be sorted in order from shallowest to deepest.

Data from multiple stations shall group data from each station together. No ordering is imposed for the station listing. A suggested practice is to sort by ascending station_id.

Sorting of data from multiple stations, times and depths shall be as follows:

- All data from a single station is grouped.
- Data from that station shall be in time order.
- All depths from that time shall be presented in order.

Conceptual illustration:

Data from Station 1 at time 1 and depth 1

Data from Station 1 at time 1 and depth 2

Data from Station 1 at time 2 and depth 1

Data from Station 1 at time 2 and depth 2
Data from Station 2 at time 1 and depth 1
Data from Station 2 at time 1 and depth 2
Data from Station 2 at time 2 and depth 1
Data from Station 2 at time 2 and depth 2

3 Conventions for Specific Observed Properties

Ed. Note: I started out planning to write a generic document for "scalars" and "vectors" at "points" and "profiles," omitting conventions for specific observed properties. However, because column order is important for quantities with multiple data and metadata values on each row, it is difficult to write a general treatment without considering specific phenomena.

In the following, the order and title of each IOOS mandatory column is specified. All mandatory columns must be included (with a null value if needed).

Any IOOS optional columns are listed next, and if used shall be placed after the mandatory columns. If any IOOS optional columns are used, they must be in the same order as specified here. If only some of the IOOS optional columns are used, then at least the first optional column and all columns up to and including the last column needed must be provided (with nulls as appropriate). A trailing set of one or more empty columns can be omitted. (In other words, do not omit columns that are not needed between columns that are needed--instead, put in a column(s) of nulls.)

Header fields shall use the column names specified below. The names are based on the CF (Climate and Forecast) Standard Names where possible.

Data providers may add additional provider-specific columns. However, if so then all the IOOS optional columns must be included (with nulls as needed) and the provider-specific columns placed after.

NOTE: In the sample responses below, individual lines of text may be wider than the page and therefore wrap onto multiple lines. Paragraph spacing in the examples has been increased slightly for clearer separation between actual data rows.

3.1 Temperature

IOOS Mandatory fields:
 sea_water_temperature (C)

NOTE: CF canonical units are K (degrees Kelvin) for temperature. IOOS uses C (degrees Celsius).

IOOS Optional fields: none

Sample response:

```

station_id:META:TEXT:61      sensor_id:META:TEXT:61      latitude [degree]
      longitude [degree]      time_ISO8601      depth [m]
      sea_water_temperature [C]

urn:ioos:station:wmo:41012:  urn:ioos:sensor:wmo:41012::watertemp1:  30.04
-80.55      2008-08-01T00:50:00Z      0.60  27.70

urn:ioos:station:wmo:41012:  urn:ioos:sensor:wmo:41012::watertemp1:  30.04
-80.55      2008-08-01T01:50:00Z      0.60  27.70

urn:ioos:station:wmo:41012:  urn:ioos:sensor:wmo:41012::watertemp1:  30.04
-80.55      2008-08-01T02:50:00Z      0.60  27.60

```

Working URL (see <http://sdf.ndbc.noaa.gov/sos/> for others):

http://sdf.ndbc.noaa.gov/sos/server.php?request=GetObservation&service=SOS&offering=urn:ioos:station:wmo:41012&observedproperty=Sea_Water_Temperature&responseformat=text/tab-separated-values

3.2 Salinity**IOOS Mandatory fields:**

sea_water_salinity (psu)

IOOS Optional fields: none

Sample response:

```

station_id:META:TEXT:61      sensor_id:META:TEXT:61      latitude [degree]
      longitude [degree]      time_ISO8601      depth [m]
      sea_water_salinity [psu]

urn:ioos:station:wmo:41012:  urn:ioos:sensor:wmo:41012::ct1:  30.04 -
80.55 2008-08-01T00:50:00Z  1.00  36.25

urn:ioos:station:wmo:41012:  urn:ioos:sensor:wmo:41012::ct1:  30.04 -
80.55 2008-08-01T01:50:00Z  1.00  36.25

urn:ioos:station:wmo:41012:  urn:ioos:sensor:wmo:41012::ct1:  30.04 -
80.55 2008-08-01T02:50:00Z  1.00  36.25

```

Working URL (see <http://sdf.ndbc.noaa.gov/sos/> for others):

http://sdf.ndbc.noaa.gov/sos/server.php?request=GetObservation&service=SOS&offering=urn:ioos:station:wmo:41012&observedproperty=Sea_Water_Salinity&responseformat=text/tab-separated-values

3.3 Sea Floor Depth (Tsunami Water Level)**IOOS Mandatory fields:**

sea_floor_depth_below_sea_surface (m)

IOOS Optional fields:

averaging interval (s)

Sample response:

```
station_id:META:TEXT:61    sensor_id:META:TEXT:61    latitude [degree]
longitude [degree]    time_ISO8601
sea_floor_depth_below_sea_surface [m]    averaging_interval [s]
urn:ioos:station:wmo:46403:    urn:ioos:sensor:wmo:46403::tsunameter0:    52.65
-156.94    2008-07-17T00:00:00Z    4509.488    900
urn:ioos:station:wmo:46403:    urn:ioos:sensor:wmo:46403::tsunameter0:    52.65
-156.94    2008-07-17T00:15:00Z    4509.464    900
urn:ioos:station:wmo:46403:    urn:ioos:sensor:wmo:46403::tsunameter0:    52.65
-156.94    2008-07-17T00:30:00Z    4509.435    900
```

Working URL (see <http://sdf.ndbc.noaa.gov/sos/> for others):

http://sdf.ndbc.noaa.gov/sos/server.php?request=GetObservation&service=SOS&offering=urn:ioos:station:wmo:46403&observedproperty=sea_floor_depth_below_sea_surface&responseformat=text/tab-separated-values

3.4 Water Surface Height (Tide Gauge Water Level)

IOOS Mandatory fields:

water_surface_height_above_reference_datum (m)
datum_id

NOTE: The datum_id is an identifier for the vertical datum to which the water-level measurements are reference. Values presently in use at IOOS are:

urn:x-noaa:def:datum:noaa::IGLD (International Great Lakes Datum)

urn:x-noaa:def:datum:noaa::MHW (mean high water)

urn:x-noaa:def:datum:noaa::MLLW (mean lower low water)

urn:x-noaa:def:datum:noaa::MSL (mean sea level)

urn:ogc:def:datum:eps::5103 (North American Vertical Datum 1988)

urn:x-noaa:def:datum:noaa::STND (station datum--values are referenced only to local station)

IOOS Optional fields: none

Sample response:

```
station_id:META:TEXT:61    sensor_id:META:TEXT:61    latitude [degree]
longitude [degree]    time_ISO8601
water_surface_height_above_reference_datum [m]    datum_id
urn:x-noaa:def:station:NOAA.NOS.CO-OPS::1617433    urn:x-
noaa:def:sensor:NOAA.NOS.CO-OPS::1617433:A1    20.03658    -155.82936
2010-03-02T13:48:00Z    0.432    urn:x-noaa:def:datum:noaa::MLLW
```

Working URL: in preparation

3.5 Winds

IOOS Mandatory fields:

wind_from_direction (degree)

wind_speed (m/s)

wind_speed_of_gust (m/s)

upward_air_velocity (m/s)

IOOS Optional fields: none

NOTE: The 'depth' value in the wind response is negative, because depth is positive below the water surface whereas the wind sensor is above the water.

Sample response:

```
station_id:META:VAR:TEXT:61      sensor_id:META:VAR:TEXT:61      latitude [degree]
      longitude [degree]      time_ISO8601      depth [m]
wind_from_direction [degree]  wind_speed [m/s]      wind_speed_of_gust [m/s]
upward_air_velocity[m/s]

urn:ioos:station:wmo:41012:      urn:ioos:sensor:wmo:41012::anemometer1:      30.04
-80.55      2008-08-01T00:50:00Z      -5.00 239.0 9.00  10.50

urn:ioos:station:wmo:41012:      urn:ioos:sensor:wmo:41012::anemometer1:      30.04
-80.55      2008-08-01T01:50:00Z      -5.00 234.0 8.30  9.30

urn:ioos:station:wmo:41012:      urn:ioos:sensor:wmo:41012::anemometer1:      30.04
-80.55      2008-08-01T02:50:00Z      -5.00 241.0 8.90  10.90
```

Working URL (see <http://sdf.ndbc.noaa.gov/sos/> for others):

<http://sdf.ndbc.noaa.gov/sos/server.php?request=GetObservation&service=SOS&offering=urn:ioos:station:wmo:41012&observedproperty=Wind&responseformat=text/tab-separated-values>

3.6 Currents

IOOS Mandatory fields:

direction_of_sea_water_velocity (degree)

sea_water_speed (cm/s)

upward_sea_water_velocity (cm/s)

NOTE: CF canonical units are m/s for the current speed quantities. IOOS uses cm/s.

IOOS Optional fields:

Error Velocity (cm/s)

platform_orientation (degree)

platform_pitch_angle (degree)

platform_roll_angle (degree)

sea_water_temperature (C)

Pct Good 3 Beam (%)

Pct Good 4 Beam (%)
 Pct Rejected (%)
 Pct Bad (%)
 Echo Intensity Beam1 (count)
 Echo Intensity Beam2 (count)
 Echo Intensity Beam3 (count)
 Echo Intensity Beam4 (count)
 Correlation Magnitude Beam1 (count)
 Correlation Magnitude Beam2 (count)
 Correlation Magnitude Beam3 (count)
 Correlation Magnitude Beam4 (count)
 Quality Flags

NOTE: The Quality Flags are packed together in a single semicolon-delimited string. The meaning of the quality flags for a particular sensor would be described in metadata (such as might be obtained using the SOS DescribeSensor operation).

NOTE: NDBC currents from MMS stations are the only dataset for which we have quality flags at present. This is a topic for further discussion, especially in the case of differing quality flags for the same phenomenon measured by different models of sensor (or post-processed by different quality-control procedures).

NOTE: In the sample dataset and sample response below, there are 9 quality flags representing the results of the following quality tests based on their position (left to right) in the flags field:

- *Flag 1 represents the overall bin status.*
- *Flag 2 represents the ADCP Built-In Test (BIT) status.*
- *Flag 3 represents the Error Velocity test status.*
- *Flag 4 represents the Percent Good test status.*
- *Flag 5 represents the Correlation Magnitude test status.*
- *Flag 6 represents the Vertical Velocity test status.*
- *Flag 7 represents the North Horizontal Velocity test status.*
- *Flag 8 represents the East Horizontal Velocity test status.*
- *Flag 9 represents the Echo Intensity test status.*

Valid flag values are:

- *0 = quality not evaluated;*
- *1 = failed quality test;*
- *2 = questionable or suspect data;*
- *3 = good data/passed quality test; and*
- *9 = missing data.*

These flag meanings and values do not necessarily apply to any other ocean currents data.

Sample response:

```

station_id:META VAR:TEXT:61    sensor_id:META VAR:TEXT:61    latitude [degree]
    longitude [degree]    time_ISO8601    bin [count] depth [m]
    direction_of_sea_water_velocity [degree]    sea_water_speed [cm/s]
  
```

```

upward_sea_water_velocity [cm/s]    error_velocity [cm/s]
platform_orientation [degree] platform_pitch_angle [degree]
platform_roll_angle [degree] sea_water_temperature [c]
pct_good_3_beam [%]    pct_good_4_beam [%]    pct_rejected [%]
pct_bad [%] echo_intensity_beam1 [count]  echo_intensity_beam2 [count]
echo_intensity_beam3 [count]  echo_intensity_beam4 [count]
correlation_magnitude_beam1 [count] correlation_magnitude_beam2 [count]
correlation_magnitude_beam3 [count] correlation_magnitude_beam4 [count]
quality_flags

urn:ioos:station:wmo:42361:    urn:ioos:sensor:wmo:42361::adcp0:    27.55 -
92.49 2010-03-02T16:03Z 1    51.00 292    16.2    -2.0    0.5
    99    99    99    205    195    215    202
    3;3;3;3;9;3;3;3;0

urn:ioos:station:wmo:42361:    urn:ioos:sensor:wmo:42361::adcp0:    27.55 -
92.49 2010-03-02T16:03Z 2    67.00 305    23.7    -0.5    -0.5
    99    99    99    194    183    198    190
    3;3;3;3;9;3;3;3;0

urn:ioos:station:wmo:42361:    urn:ioos:sensor:wmo:42361::adcp0:    27.55 -
92.49 2010-03-02T16:03Z 3    83.00 317    19.8    -0.5    -1.0
    99    99    99    202    170    216    172
    3;3;3;3;9;3;3;3;0

```

Working URL (see <http://sdf.ndbc.noaa.gov/sos/> for others):

<http://sdf.ndbc.noaa.gov/sos/server.php?request=GetObservation&service=SOS&offering=urn:ioos:station:wmo:42361&observedproperty=Curren&responseformat=text/tab-separated-values>

3.7 Waves

Waves data may be reported either using only the list of mandatory fields or by including complete spectral information in the optional fields.

The spectral information includes spectral energy, directions and polar coordinates at a number of frequencies. The number of frequencies in the spectrum shall be indicated in each data row; the remaining fields shall provide values for each spectral quantity at each frequency. Each spectral quantity shall be expressed as a list of values separated by semicolons. (That is, rather than using a separate column for each value, we pack lists of values for the same quantity as a semicolon-delimited string).

Illustrative example in the simplified case of 3 frequencies (at 0.04, 0.05 and 0.06 Hz) each of bandwidth 0.5 Hz: 3,0.04;0.05;0.06,0.005;0.005;0.005.

NOTE: The spectral quantities are packed because assigning a separate column (tab-delimited value) for each number causes the number of columns to exceed the maximum (256) supported by typical spreadsheet applications)

IOOS Mandatory fields:

sea_surface_wave_significant_height (m)
 sea_surface_wave_peak_period (s)
 sea_surface_wave_mean_period (s)
 sea_surface_swell_wave_significant_height (m)
 sea_surface_swell_wave_period (s)
 sea_surface_wind_wave_significant_height (m)
 sea_surface_wind_wave_period (s)
 sea_water_temperature (C)
 sea_surface_wave_to_direction (degree)
 sea_surface_swell_wave_to_direction (degree)
 sea_surface_wind_wave_to_direction (degree)

IOOS Optional fields:

NOTE: as discussed above, Number of Frequencies is an integer, and all the following fields are semicolon-delimited lists each containing that number of elements.

number_of_frequencies (count)
 center_frequencies (Hz)
 bandwidths (Hz)
 spectral_energy (m²/Hz)
 mean_wave_direction (degree)
 principal_wave_direction (degree)
 polar_coordinate_r1 (1)
 polar_coordinate_r2 (1)
 calculation_method
 sampling_rate

Sample response:

```

station_id:METAVAR:TEXT:61    sensor_id:METAVAR:TEXT:61    latitude [degree]
    longitude [degree]    time_ISO8601
sea_surface_wave_significant_height [m]    sea_surface_wave_peak_period [s]
sea_surface_wave_mean_period [s]
sea_surface_swell_wave_significant_height [m]
sea_surface_swell_wave_period [s]    sea_surface_wind_wave_significant_height
[m]    sea_surface_wind_wave_period [s]    sea_water_temperature [c]
sea_surface_wave_to_direction [degree]
sea_surface_swell_wave_to_direction [degree]
sea_surface_wind_wave_to_direction [degree]    number of frequencies [count]
    center frequencies [Hz] bandwidths [Hz]    spectral energy [m**2/Hz]
    mean wave direction [degree]    principal wave direction [degree]    polar
coordinate r1 [1] polar coordinate r2 [1] calculation_method
    sampling_rate [Hz]

urn:ioos:station:wmo:42002:    urn:ioos:sensor:wmo:42002::wpm1:    25.17 -
94.42 2008-12-24T04:50:00Z    2.62 7.69 5.79 0.00 0.00 2.62 7.70
    111.0    111.0 46
    0.0325;0.0375;0.0425;0.0475;0.0525;0.0575;0.0625;0.0675;0.0725;0.0775;0
.0825;0.0875;0.0925;0.1000;0.1100;0.1200;0.1300;0.1400;0.1500;0.1600;0.1700;0

```

