Marine seismology data/metadata standards

1 Introduction

1.1 Nomenclature

This document contains FDSN standards, proposed standards, and recommendations.

- Proposed standards are preceded by "**We propose**".
- Recommendations are preceded by "**We recommend**".
- FDSN standards are stated.

If an element name has a defined unit, the unit is specified as <name>.<unit>.

1.2 Adding marine-specific information to StationXML

We propose the following rule:

- Define the sub-elements that need to be specified
- Insert into StationXML as a StationXML-standardized element
- Request the addition of the information into StationXML, through FDSN WGII

1.3 StationXML-standardized element

Is an element that obeys the StationXML schema but contains information that is not specified in the schema.

Currently, StationXML-standardized elements are encoded as StationXML `<Comment>`s, with the top level in the `subject` and the sub-elements written as JSON-formatted strings.. For example, leapsecond information can be expressed as:

```
subject: “Clock correction”
content:
   leapseconds:
       list_file_entries:
           - line_text: "3692217600      37      # 1 Jan 2017"
             leap_type: '+'
```

which would be entered into the StationXML file as:

```
<Comment subject=”Clock Correction”><Value>{list_file_entries: [{line_text: '3692217600
37      # 1 Jan 2017', leap_type: '+'}]}</Value></Comment>
```

In this document, the structure is specified in the first format, for clarity.

1.4 Geophone wiring and polarity

The default polarity on short-period geophones manufactured for the exploration industry (e.g. the Geospace GS-11D) is such that a positive output voltage represents downward ground motion, in contradiction to the standard followed in passive seismology where a positive output voltage represents upward ground motion. We recommend that all ground-motion sensors follow the passive

seismology standard. For geophones, this can be readily accomplished by connecting the positive (negative) terminal of the geophone to the negative (positive) inputs to the data-logger. This can be done on the geophone itself or by switching the leads from the geophone at the data-logger inputs. Alternatively, some data-loggers allow switching polarity in software.

1.5 Clock drift corrections

There are three main possibilities for distributing data:

1. **"NOT CLOCK CORRECTED"**: No time correction applied. May be preferred by users of long-period data (>10s) because it can be easier to concatenate daily files.
2. **"CLOCK CORRECTED"**: Indicate the time correction in each record header and apply it. Allows the user to work with time-corrected but otherwise unmodified data.
3. **"RESAMPLED"**: Resample the data at the originally intended rate. Date are time-corrected and easy to concatenate/combine with other data, but could distort waveforms/spectra (needs study).

We recommend providing "CLOCK CORRECTED**"** data

1.6 Subnetwork files

We recommend using obsinfo subnetwork files to store essential information about OBS deployments.

1.7 New version of msmod

As of August 2025, additions to msmod have been written (but not tested), allowing direct clock drift corrections for piecewise linear, polynomial, or cubic spline estimates of the clock drift.

A further, requested modification, should allow leap seconds to be inserted in miniseed files using a single call. If this modification is made, the call recommended in the Positive leap second and Negative leap second sections would simplify to

```
msmod --lsp '2016,183,00:00:00'
```

and

```
msmod --lsn '2016,183,00:00:00'
```

respectively

1.8 Saving processing information

We recommend that processing done on data files (from data download to delivery to the data center) should be recorded in text-based, structured files. The JSON process-steps format is an example.

1.9 Station names for repeated deployments

The IASPEI Station Coding Standard recommends that station and/or location codes be changed if the associated sensors are moved far enough to result in a significant *teleseismic* travel-time residual

discrepancy. We recommend the same, except that the basis for what is a significant travel-time residual discrepancy should depend on your study.

If you change the station name between deployments, we recommend incrementing the last *N* characters of the station name, e.g. ``STAA``, ``STAB``, ``STAC``, or ``STA01``, ``STA02``, ``STA03``, etc. The value of *N* depends on the maximum number of deployments you will possibly make and the characters you use in the incrementor (numeric, alphabetic, or alphanumeric).

2 Source identifiers

The following are with respect to FDSN Source Identifiers, documented at https://docs.fdsn.org/projects/source-identifiers/en/v1.0/channel-codes.html#source-and-subsource-codes.

2.1 Subsource codes

2.1.1 Seismometer channels

N,E,Z are as defined subsource codes in the FDSN Source Identifiers document

1,2,3 are used for orthogonal components that do not match the N,E,Z or U,V,Z dip/azimuth values. The metadata must be consulted for the orientation of channels with subsource code 1, 2, and/or 3.

We recommend that, if there are two orthogonal horizontal channels that are not known to be aligned to N and E, they be named 1 and 2, with 2 oriented 90° clockwise of 1 when viewed from above.

The azimuths should be expressed in the metadata as:

- 1: **<Azimuth unit="DEGREES" plusError=180 minusError=180> 0 </Azimuth>**
- **2: <Azimuth unit="DEGREES" plusError=180 minusError=180> 90 </Azimuth>**

If the vertical channel is inverted (NOT RECOMMENDED, see section 1.4), **we recommend** using subsource code 3 and <Dip unit=”DEGREES”> 90 </Dip>

If the three components are orthogonal and there are not two horizontals, **we recommend** using subsource codes 1, 2 and 3. The relative orientation between the channels should be provided in the metadata, with plusError=180 and minusError=180 for all Azimuths and Dips.

2.1.2 Pressure channels

Set Dip=90. if the recorded value DECREASES for a pressure increase.

Set Dip=-90. if the value INCREASES for a pressure increase.

This gives the same polarity as a vertical channel for UPGOING waves (first arrivals).

We remind that there are different subsource codes for hydrophones (H), differential pressure gauges (G) and absolute pressure gauges (O).

3 StationXML metadata

The following recommendations and reminders are with respect to StationXML 1.2, documented at https://docs.fdsn.org/projects/stationxml.

3.1 Clock drift

We recommend specifying clock drifts using UTC datetimes. The datetimes should be in ISO8601 format, followed by a "Z" to unambiguously specify UTC.

We propose the following format:

Subject: “Clock Correction”

Contents:

```
drift:
    type: 'piecewise_linear'  # or 'cubic_spline' or 'polynomial {a0 a1 a2 ...}'
    instrument: 'Seascan MCXO'
    instrument_nomimal_drift_rate: 1e-8
    reference: 'GPS'
    syncs_instrument_reference:
        - ["2016-09-10T00:00:00Z", "2016-09-10T00:00:00Z"]
        - ["2017-01-12T00:00:01Z", "2017-01-12T00:00:00.415Z"]
        - ["2017-07-13T11:25:01Z", "2017-07-13T11:25:00.6189Z"]
```

which is written as a Station-level StationXML-standardized Element.:

```
<Comment subject=”Clock Correction”><Value>“{drift: {instrument: Seascan MCXO,
 instrument_nominal_drift_rate: 1e-8, reference: GPS, type: piecewise_linear,
 syncs_instrument_reference: [['2016-09-10T00:00:00Z', '2016-09-10T00:00:00Z'],
 ['2017-01-12T00:00:01Z', '2017-01-12T00:00:00.415Z'], ['2017-07-13T11:25:01Z',
 '2017-07-13T11:25:00.6189Z',]]}}”</Value></Comment>
```

The `time_base`, `nominal_drift_rate` and `reference` fields are optional. In the simplest case of a synchronization at the start and end of an experiment, and assuming purely linear drift, there would only be two items in syncs_instrument_reference.

If no drift is measured, the StationXML-standardized Element should be:

```
<Comment subject=”Clock Correction”><Value></Value></Comment>
```

meaning that we know that there is a clock drift but we don’t know what it is.

3.2 Leap seconds

Specified using information from the `leap-seconds.list` file, available online at several sites, including https://data.iana.org/time-zones/tzdb/leap-seconds.list. The user should verify that the "File expires on" date is later than the last instrument channel's end-date.

We propose the following format:

Subject: “Clock Correction”

Contents:

```
leapseconds:
```

```
list_file_entries:
    - line_text: "3692217600      37      # 1 Jan 2017"
      leap_type: '+'
```

to be written as a Network-level StationXML-standardized Element.:

```
<Comment subject=”Clock Correction”><Value>{leapseconds: {list_file_entries: [{
 line_text: '3692217600      37      # 1 Jan 2017', leap_type:
'+'}]}}</Value></Comment>
```

where

- `list_file_entries` is an array/list, to allow for more than one leap-second during a deployment.
- `line_text` should be directly copied from the `leap-seconds.list` file
- `leap_type` indicates whether the 2nd number in the list_file line_text ("37" in the above example) is greater than the previous line's value ("+") or less than the previous line's value ("-"). Up to August 2025, all leap seconds have been type "+"

In general, the NOT_CLOCK_CORRECTED data will not have a leap-second correction, and this correction should not be applied to this data. If future dataloggers integrate leap-second correction in their raw data, this should be specified by adding the following Station-level StationXML-standardized Element:

```
<Comment subject=”Clock Correction”><Value>{leapseconds:
{leapsecond_is_integrated_into_not_clock_corrected_miniseed: True}}</Value></Comment>
```

3.3 Specifying drift in data that contains a leap second

- **We propose** that the instrument clock drift times provided in the stationXML file be leap-second corrected

3.4 Data completeness

We recommend using Station `<StartDate>` and `<EndDate>` to specify when the data was supposed to start and end, and Channel `<StartDate>` and `<EndDate>` to specify when it actually starts and ends for each channel.

We recommend keeping all of the recorded data, including "noisy" or "bad".

3.5 Leveling system

The instrument's leveling system can affect the quality and absolute values of measurables.

We propose implementing this as an "Equipment" element at the appropriate (Station or Channel) level, with `<Type>Leveler</Type>` and `<Description>` a JSON-encoded string of the following elements.

Name	Type	Units	Required?	Description
`threshold.deg`	float	arc-degrees	Yes	deviation from vertical (degrees) which will trigger releveling

accuracy.deg	float	arc-degrees	Yes	Maximum deviation from vertical accepted after relevel
max_relevel_interval.h	float	hours	No	Longest interval between level checks during the deployment
n_relevels	Int		No	Number of relevels performed during the deployment
relevels	list	list	No	List of [date, level_before.deg, level_after.deg] for each relevel
description	str		Yes	Description of the leveling system (for example, “gravity-based gimbal system in oil”, or “Guralp Aquarius automatic compensation system”). Manufacturer-defined description should be used if available

For example:

```
<Equipment>
    <Type>Leveler</Type>
    <Description>{threshold.deg: 2.5, accuracy.deg: 0.25, max_relevel_interval.h: 168,
n_relevels: 12, description: ‘SIO-Nanometrics BBOBS leveling system’}</Description>
    <Manufacturer>Scripps Inst. Oceanography</Manufacturer>
    <SerialNumber>FR02</SerialNumber>
</Equipment>
```

We will request, for a future version of StationXML, a new superelement of the ``Equipment`` element, named “Leveler” and with the subelements currently specified in <Description> added.

3.6 Positions and Orientations

We recommend using the `plusError`, `minusError` and `measurementMethod` attributes to specify uncertainties in Latitude, Longitude and Elevation and how these values were measured/estimated.

We recommend using the `plusError`, `minusError` and `measurementMethod` attributes to specify uncertainties in estimated Dip and Azimuth values and how these values were measured/estimated.

3.7 Deployments in lakes

We remind that the `<WaterLevel>` elements should be set to the elevation of the lake surface

3.8 Standard values that not all seismologists may know

Within each `<Channel>`, set `<Type>CONTINUOUS</Type>` and `<Type>GEOPHYSICAL</Type>`

For pressure sensor channel responses, `InputUnits` should be `Pa`, without a `Description`, i.e.:

```
<InputUnits><Name>Pa</Name></InputUnits>
```

4 miniSEED data

The following recommendations and reminders are with respect to miniSEED2, documented at https://www.fdsn.org/pdf/SEEDManual_V2.4.pdf , and miniSEED3, documented at https://fdsn-docs.readthedocs.io/projects/miniseed3/en/latest/index.html .

4.1 Clock drift correction

For "NOT CLOCK CORRECTED" data.

1. **We propose,** for Indicating:
    a. miniSEED2: Set the data quality flag to"D"
    b. miniSEED3: Set root -> FDSN -> DataQuality to “D”
2. **We recommend,** for Creating:
    a. miniSEED2: Put time correction in record header field 16 and set field 12 bit 1 to 0.
    b. miniSEED3: No specific header information

For "CLOCK CORRECTED" data.

1. **We propose,** for Indicating:
    a. miniSEED2: Set the data quality flag to "Q"
    b. miniSEED3: Set root -> FDSN -> DataQuality to “Q”
1. **We recommend,** for Creating:
    a. Calculate a new time drift for each record
    b. miniSEED2:
        i. In record header field 16 (“Time Correction”), specify correction added to original time (in units of 0.0001 seconds).
        ii. Set record header field 12, bit 1 (“Activity Flag, time correction applied”) to 1.
    c. miniSEED3:
        i. In root -> FDSN -> Time -> Correction, specify correction added to original time (text, in units of float seconds)

For "RESAMPLED" data.

1. **We recommend** for Indicating:
    a. miniSEED 2/3: Define another channel code (modified data)
2. No recommendation for Creating

4.1.1 Unmeasured clock drift

If clock drift is expected, but there is no measure/estimate of it, **we recommend**:

- miniSEED2:
    - Provide data as "D". set bit 7 of data quality flag ("time tag is questionable") to 1.
    - Add blockette 500, field 10 ("Clock status") indicating that there is an unmeasured drift (for example: "Unmeasured clock drift on Seascan MCXO, expected order = 1e-8")
- miniSEED3
    - No recommendation yet.

4.1.2 Providing both CLOCK CORRECTED and NOT CLOCK CORRECTED data

Most FDSN-compatible data centers store data in SeisComP Data Structure (SDS), which does not distinguish between data with different data qualities. If you want to provide both **CLOCK CORRECTED** and **NOT CLOCK CORRECTED** data, they must be in the same files.

4.2 Leap second correction

Leap seconds should be corrected in **CLOCK CORRECTED** data and the record containing the leap second should be flagged. Note there there have only been positive leap seconds to date (as of August 2025).

4.2.1 Positive leap second (61 seconds in the minute)

The procedure is to:

- Modify the miniSEED files:
    - Shift all record times AFTER the leap second back one second.
    - Set activity flag bit 4 to 1 in the header of the record containing the leap second.
- Change `end_sync_instrument` to be one second earlier than what the instrument indicated

Here is an example of modifiying the miniSEED files for a positive leap-second at 23:59:60 on day 182, 2016:

```
msmod --timeshift -1 -ts 2016,182,23:59:59.999999'
msmod –-actflags ‘4,1’ –tsc 2016,182,23:59:59.999999 –tec 2016,182,23:59:59.999999
```

4.2.2 Negative leap second (59 seconds in the minute)

The procedure is to:

- Modify the miniSEED files
    - Shift all record times AFTER the leap second forward one second.
    - Set activity flag bit 5 to 1 in the header of the record containing the leap second.
- Change `end_sync_instrument` to be one second later than what the instrument indicated

4.3 Correcting drift in data that contains a leap second

- **We propose** that “NOT CLOCK CORRECTED” data not be leap-second corrected, and that “CLOCK CORRECTED” data be leap-second corrected
- **We propose** that the drift correction be performed in two steps:
    - Applying the drift correction (section)
    - Then applying the leap second correction (section)

Note: If you use the data to calculate the drift rate, first generate LEAP-CORRECTED data from the NOT CLOCK CORRECTED DATA. Use this data to calculate the clock drift. Then throw the LEAP-CORRECTED data away.