Putting ocean bottom seismometer data in Résif

Wayne C. Crawford, Olivier Dewée, Rima Bouazzouz, David Wolyniec

Putting ocean bottom seismometer data in RESIF

Wayne Crawford[1] (crawford@ipgp.fr), Olivier Dewee[2], Rima Bouazzouz[3] and David Wolyniec[3]

1: Institut de Physique du Globe de Paris, Marine Geosciences Team, 1 rue Jussieu, 75238 Paris Cedex 5, France
2: INSU Division Technique, UPS 855, 1 Place Aristide Briand, 92195 Meudon Cedex, France
3: RESIF Data center, ISTERRE, Université Grenoble Alpes, CS40700, 38058 Grenoble Cedex 9

Overview

Few ocean bottom seismometer experiments are currently available on EIDA data centers, in part because of the difficulty of preparing their data and metadata for data centers, including:
- Non-standard dataloggers which are unreferenced in the Nominal Response Library (NRL) and which have non-standard data formats
- The inability to synchronize the sensor clock with GPS while at the seafloor
- A lack of standards for OBS-specific parameters

We present a 3-part, harmonized approach to simplify data/metadata preparation and standardize the output using FAIR principles:
1) **Standards for OBS data and metadata**
2) Software to **prepare OBS datafiles for EIDA**
3) Software and «information files» to **prepare metadata for EIDA**

1) Standards

Standards have been proposed through FDSN Working Group 5 (Mobile Instrumentation) and SERA Deliverable 4.2. The principal standards are:
- **SEED «Data quality» codes:**
  - **'D'** : no time correction applied
  - '**Q**': time correction applied and fully specified in the miniSEED header.
- **Channel naming**:
  - Horizontal channels should be called '1' and '2', with '1' corresponding geometrically to 'N'. and '2' to 'W'
  - Geophone vertical channels (positive downward) should be named '3'
- **Channel azimuths/dips** in StationXML:
  - **'1'** has azimuth = 0.0, uncertainty = 180.0
  - **'2'** has azimuth = 90.0, uncertainty = 180.0
  - **'Z'** has dip = -90.0
  - **'3'** has dip = 90.0
  - **'H'** (hydrophone) has dip=
    - -90.0 if a postive voltage corresponds to a positive pressure
    - 90.0 if a positive voltage corresponds to a negative pressure.
- **Comments**
  - Processing steps (clock correction, leap seconds...) should be specified in comments. Clock corrections should be specified as absolute dates, not «slew» or «drift» as the polarity of these latter are ambiguous.

2) Preparing data: SDPCHAIN

Programs to transform obs data into time-corrected SeisComp3 Data Structure (SDS), while creating a Provenance file. The software works on uncorrected miniSEED files (considered the primary «standard» output of OBS parks).

mscat: concatenate miniSEED files
msdrift: apply a clock drift correction
msleap: add a leap-second correction
ms2sds: convert miniSEED file(s) to SeisComp3 Data Structure
sdp-process: run a Unix command-line program and appends the provenance file.

This package is not yet ready for distribution, but interested parties can obtain it and help us evaluate and upgrade it.

Example SDPCHAIN Provenance file (JSON format)


```
{"steps":[
    {"application":{"name":"lcFix.py","version":"0.66","description":"Fix common bugs in LCHEAPO data files"},
    "execution":{
        "commandline":"\/opt\/lcheapo\/bin\/lcfix MOCE.raw.lch -d \/Volumes\/Wayne_Data\/_NoBackup\/MOCE -i . -o lcheapo_fixed",
        "date":"2019-05-07T09:00:35",
        "messages":["  MOCE.raw.lch=>\/Volumes\/Wayne_Data\/_NoBackup\/MOCE\/lcheapo_fixed\/MOCE.fix.lch: Finished at block 3988993 (0 BUG1s, 0 BUG2s, 0 Time Tears, 0 unexpected header values)"],
        "parameters":{
            "base_directory":"\/Volumes\/Wayne_Data\/_NoBackup\/MOCE","dryrun":false,"forceTime":false,
            "input_directory":".","input_files":["MOCE.raw.lch"],"output_directory":"lcheapo_fixed",
            "output_files":["\/Volumes\/Wayne_Data\/_NoBackup\/MOCE\/lcheapo_fixed\/MOCE.fix.lch"],"verbosity":0},
            "return_code":0,"tools":[]}},
    {"application":{"name":"lc2ms","comment":"Transform LCHEAPO data files to miniSEED","version":"1.2"},
    "execution":{
        "commandline":"lc2ms MOCE.fix.lch -d \/Volumes\/Wayne_Data\/_NoBackup\/MOCE -i lcheapo_fixed -o miniseed_basic -m :%E.%S.00.%C.%Y.%D.%T.mseed --experiment 1T --sitename MOCE --obstype SPOBS2 --sernum 17 -p \/opt\/lcheapo\/config\/lc2ms.properties",
        "date":"2019-05-07T12:01:21",
        "messages":[
             "L-CHEAPO file processed : \/Volumes\/Wayne_Data\/_NoBackup\/MOCE\/lcheapo_fixed\/MOCE.fix.lch",
             "Mini-SEED file generated : \/Volumes\/Wayne_Data\/_NoBackup\/MOCE\/miniseed_basic\/1T.MOCE.00.EDH.2019.085.135759.mseed",
             "Mini-SEED file generated : \/Volumes\/Wayne_Data\/_NoBackup\/MOCE\/miniseed_basic\/1T.MOCE.00.EH1.2019.085.135759.mseed",
             "Mini-SEED file generated : \/Volumes\/Wayne_Data\/_NoBackup\/MOCE\/miniseed_basic\/1T.MOCE.00.EH2.2019.085.135759.mseed",
             "Mini-SEED file generated : \/Volumes\/Wayne_Data\/_NoBackup\/MOCE\/miniseed_basic\/1T.MOCE.00.EH3.2019.085.135759.mseed"],
        "parameters":{
            "experiment":"1T",
            "io_files":[{
                "input_file":"\/Volumes\/Wayne_Data\/_NoBackup\/MOCE\/lcheapo_fixed\/MOCE.fix.lch",
                "output_files":[
                    "\/Volumes\/Wayne_Data\/_NoBackup\/MOCE\/miniseed_basic\/1T.MOCE.00.EDH.2019.085.135759.mseed",
                    "\/Volumes\/Wayne_Data\/_NoBackup\/MOCE\/miniseed_basic\/1T.MOCE.00.EH1.2019.085.135759.mseed",
                    "\/Volumes\/Wayne_Data\/_NoBackup\/MOCE\/miniseed_basic\/1T.MOCE.00.EH2.2019.085.135759.mseed",
                    "\/Volumes\/Wayne_Data\/_NoBackup\/MOCE\/miniseed_basic\/1T.MOCE.00.EH3.2019.085.135759.mseed"]
            }],
            "obs":{"device_type":"SPOBS2","position_name":"MOCE","serial_number":"17"},
            "seed":{"byte_order":1,"encoding_format":10,"record_length":4096}},
        "return_code":0,
        "tools":[{
            "command_parameters":"\/Volumes\/Wayne_Data\/_NoBackup\/MOCE\/lcheapo_fixed\/MOCE.fix.lch -o T-1557219681622 -p MOCE -s 17 -t SPOBS2 -c \/opt\/lcheapo\/config\/obs-config.csv -r 4096 -e 10 -b 1",
            "name":"lch2mseed","exit_status":0,"version":"lch2mseed version: 0.2\n"}]}},
    {"application":{"name":"sdp-process","description":"Runs a command-line tool and saves information to process-steps.json","version":"2.0"},
    "execution":{
        "commandline":"[\"\/opt\/sdpchain\/bin\/sdp-process\", \"-d\", \"\/Volumes\/Wayne_Data\/_NoBackup\/MOCE\", \"-c=Forcing data quality to D\", \"--cmd=msmod --quality D -i miniseed_basic\/*.mseed\"]",
        "comment":"Forcing data quality to D","date":"2019-05-07T09:01:51","return_code":0,
        "parameters":{
            "base_dir":"\/Volumes\/Wayne_Data\/_NoBackup\/MOCE","exec_comment":"Forcing data quality to D",
            "tool_cmd_line":"msmod --quality D -i miniseed_basic\/*.mseed"},
        "tools":[{
            "name":"msmod","command_parameters":"--quality D -i miniseed_basic\/*.mseed",
            "description":"msmod - Modify Mini-SEED data: 1.2","version":"msmod version: 1.2"}]}},
    {"application":{"name":"ms2sds","comment":"Transform miniSEED files to SeisComp3 Data Structure.\n","version":"1.0"},
    "execution":{
        "commandline":"ms2sds 1T.MOCE.00.EDH.2019.085.135759.mseed 1T.MOCE.00.EH1.2019.085.135759.mseed 1T.MOCE.00.EH2.2019.085.135759.mseed 1T.MOCE.00.EH3.2019.085.135759.mseed -d
\/Volumes\/Wayne_Data\/_NoBackup\/MOCE -i miniseed_basic -o SDS_uncorrected --network 1T --station MOCE -a SDS -p \/opt\/sdpchain\/config\/ms2sds.properties",
        "comment":"","date":"2019-05-07T12:01:52",
        "messages":[
            "MSEED files processed for location\/channel (00\/EDH)",
            "MSEED files processed for location\/channel (00\/EH1)",
            "MSEED files processed for location\/channel (00\/EH2)",
            "MSEED files processed for location\/channel (00\/EH3)"
        ],
        "parameters":{
            "input_files":[
                "\/Volumes\/Wayne_Data\/_NoBackup\/MOCE\/miniseed_basic\/1T.MOCE.00.EDH.2019.085.135759.mseed",
                "\/Volumes\/Wayne_Data\/_NoBackup\/MOCE\/miniseed_basic\/1T.MOCE.00.EH2.2019.085.135759.mseed",
                "\/Volumes\/Wayne_Data\/_NoBackup\/MOCE\/miniseed_basic\/1T.MOCE.00.EH1.2019.085.135759.mseed",
```


Step 1

Step 2

Step 3

Step 4

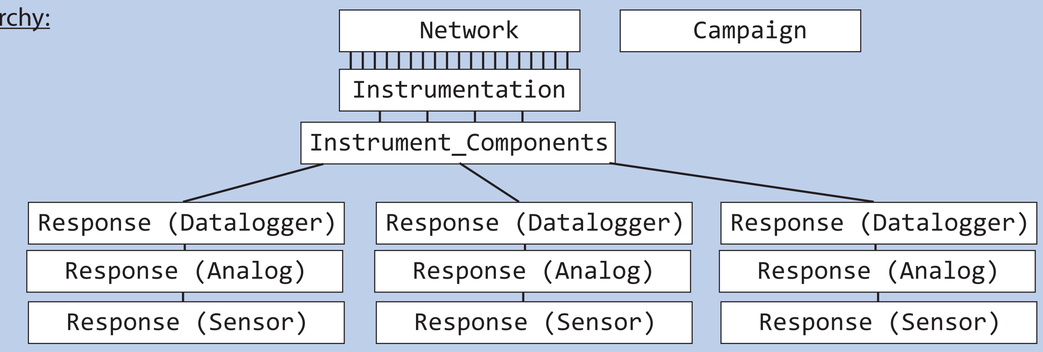
3) Preparing metadata: obsinfo

obsinfo is a specification for «information files» and an associated Open Source software library that can be used to prepare StationXML metadata and data preparation scripts.

Information files are divided into two categories:
Instrumentation-level: describe the instruments, components and responses in a hierarchical format that avoids repetition. These can be prepared by the OBS facility well before their missions, and updated as equipment is added or modified
Campaign-level: describes an individual measurement campaign.

All of the files are in YAML format (JSON is also accepted).

Hierarchy:



Software. Open-source python software (available on GitHub and PyPI) to process the information files. Routines can be called, and several command-line programs are available:
- obsinfo-validate: validates an individual information file
- obsinfo-print: prints information from an information file, often validating its links to files «downstream»
- obsinfo-makeSTATIONXML: creates StationXML files from a Network file
- obsinfo-make_SDPCHAIN_scripts: creates a processing chain to drift correct miniSEED data, correct leap-seconds, and put into SeisComp3 Data Structure (SDS)
- obsinfo-make_LCHEAPO_scripts makes a script using software specific to the INSU/IPGP OBS park to convert their data to miniseed format. Other parks can write similar code for their data.

Example obsinfo Network file (YAML format)


```
---
format_version: "0.107"
revision:
    authors:
        -
            first_name : "Wayne"
            last_name : "Crawford"
            email : "crawford@ipgp.fr"
    date: "2019-04-08"
network:
    facility:
        ref_name: "INSU-IPGP"
        full_name: "INSU-IPGP OBS Park"
    campaign_ref_name: "SMARTIES"
    network_info:
        code: "XX"
        start_date: "2019-07-18"
        end_date: "2019-08-13"
        description: "Seismic study of Romanche Fracture Zone"
        comments: []
    ref_file:{$ref: "../instrumentation.yaml"}
    ref_file_type: "instrumentation"
    stations:
        "SM19":
            site: "Fracture zone"
            start_date: "2019-07-18T12:00:00Z"
            end_date: "2019-08-13T03:45:00Z"
            location_code: "00"
            instruments:
                -
                    ref_code: "SPOBS2"
                    serial_number: "01"
                    channel_mods:
                        base: {datalogger: {config: "125sps"}}
            locations:
                "00":
                    position: {lat: -0.22987, lon: -16.32343, elev: -3724}
                    depth.m: 0
                    geology: "unknown"
                    vault: "Sea floor"
                    uncertainties.m: {lat: 200, lon: 200, elev: 20}
                    localisation_method: "Surface release"
            processing:
                - clock_correct_linear_drift:
                    time_base: "Seascan MCXO, ~1e-8 nominal drift"
                    reference: "GPS"
                    start_sync_instrument: 0
                    start_sync_reference: "2019-07-18T07:52:00"
                    end_sync_instrument: "2019-08-13T03:42:10"
                    end_sync_reference: "2019-08-13T03:42:09.9989"
        "SM18":
            site: "Fracture zone"
            start_date: "2019-07-18T13:00:00"
            end_date: "2019-08-13T03:45:00"
            station_location: "00"
            instruments:
                -
                    ref_code: "SPOBS2"
                    serial_number: "03"
                    channel_mods:
                        base: {datalogger: {config: "125sps"}}
            locations:
                "00":
                    position: {lat: -0.03483, lon: -16.68130, elev: -3729}
                    depth.m: 0
                    geology: "unknown"
                    vault: "Sea floor"
                    uncertainties.m: {lat: 200, lon: 200, elev: 20}
                    localisation_method: "Surface release"
            processing:
                - clock_correct_linear_drift:
                    time_base: "Seascan MCXO, ~1e-8 nominal drift"
                    reference: "GPS"
                    start_sync_instrument: 0
                    start_sync_reference: "2019-07-19T09:34:00"
                    end_sync_instrument: "2019-08-14T05:43:10"
                    end_sync_reference: "2019-08-14T05:43:09.8796"
```


Provenance

Network parameters

Reference file

StationXML-related

Processing

First Station

StationXML-related

Processing

Second Station

Where to find the files:

Standards:
SERA deliverable : http://www.sera-eu.org/en/Dissemination/deliverables/ , then D4.2
Latest proposal : http://www.ipgp.fr/~crawford/Documents/OBS_data_standards.pdf
obsinfo: https://pypi.org/project/obsinfo/ and https://github.com/WayneCrawford/obsinfo
SDPCHAIN: write to crawford@ipgp.fr or olivier.dewee@cnrs.fr