EarthScopeOceans

Data Committee Report and Recommendations

Presented by Tim Ahern

ESO Data Committee Working Group

- Tim Ahern, IRIS emeritus USA
- Joel Simon, Princeton USA
- Christoph Waldman, MARUM Germany

Results of November ESO Data Committee Meeting

- Progress in Data Flow and Access
- GeoCSV format for capturing rapidly changing metadata
- Status of ESO efforts in the FDSN
- Seismic data as an Essential Ocean Variable (EOV)
- Exchange of data with CTBTO
- Building the ESO Data System
  - ESO Data Centre Requirements
  - Recommended ESO Data Release Policy

Progress in Data Flow and Access

- ESO data from French and Princeton Mermaids have been available for several months.
- At the time of this meeting latest data are from November 25, 2021
- Joel has implanted a workflow since the Data Committee meeting to reduce the latency.
- 3 Princeton and 1 French MERMAID data are available without delay
  - Other stations available to authenticated users
  - All data available 2 years after real time using a rolling release window

GeoCSV Format

- A CSV format to capture metadata that changes slowly in time is operational
  - Mostly for positional data of MERMAIDS as well as a few other parameters

Example of newer GEOCSV format

#dataset: GeoCSV 2.0													
#created: 2021-11-15T20:41:12Z													
#automaid: v3.5.0 (https://github.com/earthscopeoceans/automaid [doi: 10.5281/zenodo.5057096])													
#delimiter: ','													
#lineterminator: '\n'													
#field_unit	ISO_8601	unitless	unitless	unitless	unitless	degrees_north	degrees_east	meters	meters	unitless	hertz	seconds	seconds
#field_type	datetime	string	string	string	string	float	float	float	float	string	float	float	float
MethodIdentifier	StartTime	Network	Station	Location	Channel	Latitude	Longitude	Elevation	Depth	SensorDescription	SampleRate	TimeDelay	TimeCorrection
Measurement:GPS:u-blox_NEO-M8N	2018-08-05T13:23:15Z	MH	P0008	nan	nan	-12.008233	-172.023102	0	0	MERMAIDHydrophone(452.020)	nan	0.000030	nan
Measurement:GPS:u-blox_NEO-M8N	2018-08-05T13:32:46Z	MH	P0008	nan	nan	-12.006967	-172.018723	0	0	MERMAIDHydrophone(452.020)	nan	-0.000062	nan
Measurement:GPS:u-blox_NEO-M8N	2018-08-06T13:47:22Z	MH	P0008	nan	nan	-12.047700	-172.013565	0	0	MERMAIDHydrophone(452.020)	nan	0.437377	nan
Measurement:GPS:u-blox_NEO-M8N	2018-08-06T13:48:40Z	MH	P0008	nan	nan	-12.047684	-172.013687	0	0	MERMAIDHydrophone(452.020)	nan	0.000000	nan
Measurement:GPS:u-blox_NEO-M8N	2018-08-06T13:54:20Z	MH	P0008	nan	nan	-12.047584	-172.014252	0	0	MERMAIDHydrophone(452.020)	nan	0.000000	nan
Algorithm:automaid:v3.5.0	2018-08-08T01:42:00Z	MH	P0008	0	BDH	-12.074427	-171.996506	0	1531	MERMAIDHydrophone(452.020)	20.0	nan	-0.290870
Measurement:GPS:u-blox_NEO-M8N	2018-08-15T00:10:52Z	MH	P0008	nan	nan	-12.205566	-171.903763	0	0	MERMAIDHydrophone(452.020)	nan	1.643707	nan
Measurement:GPS:u-blox_NEO-M8N	2018-08-15T00:12:58Z	MH	P0008	nan	nan	-12.205900	-171.904053	0	0	MERMAIDHydrophone(452.020)	nan	0.000000	nan
Measurement:GPS:u-blox_NEO-M8N	2018-08-15T00:23:25Z	MH	P0008	nan	nan	-12.207367	-171.905502	0	0	MERMAIDHydrophone(452.020)	nan	-0.000184	nan
Measurement:GPS:u-blox_NEO-M8N	2018-08-15T00:26:41Z	MH	P0008	nan	nan	-12.207784	-171.905991	0	0	MERMAIDHydrophone(452.020)	nan	0.000000	nan
Algorithm:automaid:v3.5.0	2018-08-16T05:47:33Z	MH	P0008	0	BDH	-12.231257	-171.890152	0	1527	MERMAIDHydrophone(452.020)	20.0	nan	-0.242693
Algorithm:automaid:v3.5.0	2018-08-17T15:43:28Z	MH	P0008	0	BDH	-12.255579	-171.868423	0	1521	MERMAIDHydrophone(452.020)	20.0	nan	-0.523292
Measurement:GPS:u-blox_NEO-M8N	2018-08-17T21:17:47Z	MH	P0008	nan	nan	-12.261200	-171.865814	0	0	MERMAIDHydrophone(452.020)	nan	0.569366	nan
Measurement:GPS:u-blox_NEO-M8N	2018-08-17T21:20:36Z	MH	P0008	nan	nan	-12.261483	-171.865891	0	0	MERMAIDHydrophone(452.020)	nan	0.000000	nan
Measurement:GPS:u-blox_NEO-M8N	2018-08-17T21:30:53Z	MH	P0008	nan	nan	-12.262450	-171.866318	0	0	MERMAIDHydrophone(452.020)	nan	-0.000153	nan
Measurement:GPS:u-blox_NEO-M8N	2018-08-17T21:34:04Z	MH	P0008	nan	nan	-12.262750	-171.866470	0	0	MERMAIDHydrophone(452.020)	nan	0.000000	nan
Algorithm:automaid:v3.5.0	2018-08-19T00:20:02Z	MH	P0008	0	BDH	-12.279121	-171.850571	0	1514	MERMAIDHydrophone(452.020)	20.0	nan	-0.212875
Algorithm:automaid:v3.5.0	2018-08-19T00:39:30Z	MH	P0008	0	BDH	-12.279303	-171.850372	0	1509	MERMAIDHydrophone(452.020)	20.0	nan	-0.215457
Algorithm:automaid:v3.5.0	2018-08-19T00:52:59Z	MH	P0008	0	BDH	-12.279428	-171.850235	0	1511	MERMAIDHydrophone(452.020)	20.0	nan	-0.217243
Algorithm:automaid:v3.5.0	2018-08-19T02:19:21Z	MH	P0008	0	BDH	-12.280232	-171.849335	0	1515	MERMAIDHydrophone(452.020)	20.0	nan	-0.228692
Algorithm:automaid:v3.5.0	2018-08-19T03:07:24Z	MH	P0008	0	BDH	-12.280680	-171.848831	0	1515	MERMAIDHydrophone(452.020)	20.0	nan	-0.235061
Algorithm:automaid:v3.5.0	2018-08-19T04:20:04Z	MH	P0008	0	BDH	-12.281357	-171.848068	0	1519	MERMAIDHydrophone(452.020)	20.0	nan	-0.244693
Algorithm:automaid:v3.5.0	2018-08-19T04:29:09Z	MH	P0008	0	BDH	-12.281442	-171.847977	0	1519	MERMAIDHydrophone(452.020)	20.0	nan	-0.245898
Algorithm:automaid:v3.5.0	2018-08-19T05:59:02Z	MH	P0008	0	BDH	-12.282278	-171.847031	0	1519	MERMAIDHydrophone(452.020)	20.0	nan	-0.257811
Algorithm:automaid:v3.5.0	2018-08-19T06:37:18Z	MH	P0008	0	BDH	-12.282635	-171.846634	0	1510	MERMAIDHydrophone(452.020)	20.0	nan	-0.262883
Algorithm:automaid:v3.5.0	2018-08-19T07:07:36Z	MH	P0008	0	BDH	-12.282917	-171.846313	0	1520	MERMAIDHydrophone(452.020)	20.0	nan	-0.266900

Updated GeoCSV file to enable full reproducibility of results and express depth in dbar not meters

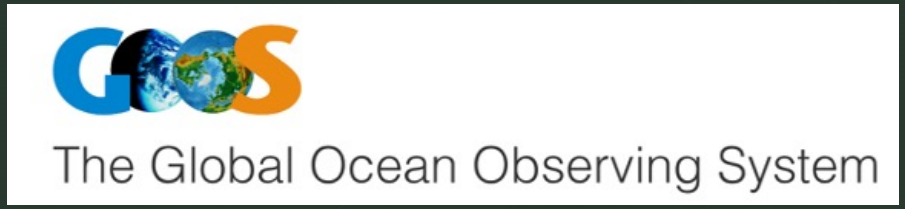
#dataset: GeoCSV 2.0													
#created: 2021-11-15T20:41:12Z													
#automaid: v3.5.0 (https://github.com/earthscopeoceans/automaid [doi: 10.5281/zenodo.5057096])													
#delimiter: ','													
#lineterminator: '\n'													
#field_unit	ISO_8601	unitless	unitless	unitless	unitless	degrees_north	degrees_east	meters	meters	unitless	hertz	seconds	seconds
#field_type	datetime	string	string	string	string	float	float	float	float	string	float	float	float
MethodIdentifier	StartTime	Network	Station	Location	Channel	Latitude	Longitude	Elevation	Depth	SensorDescription	SampleRate	TimeDelay	TimeCorrection
Measurement:GPS:u-blox_NEO-M8N	2018-08-05T13:23:15Z	MH	P0008	nan	nan	-12.008233	-172.023102	0	0	MERMAIDHydrophone(452.020)	nan	0.000030	nan
Measurement:GPS:u-blox_NEO-M8N	2018-08-05T13:32:46Z	MH	P0008	nan	nan	-12.006967	-172.018723	0	0	MERMAIDHydrophone(452.020)	nan	-0.000062	nan
Measurement:GPS:u-blox_NEO-M8N	2018-08-06T13:47:22Z	MH	P0008	nan	nan	-12.047700	-172.013565	0	0	MERMAIDHydrophone(452.020)	nan	0.437377	nan
Measurement:GPS:u-blox_NEO-M8N	2018-08-06T13:48:40Z	MH	P0008	nan	nan	-12.047684	-172.013687	0	0	MERMAIDHydrophone(452.020)	nan	0.000000	nan
Measurement:GPS:u-blox_NEO-M8N	2018-08-06T13:54:20Z	MH	P0008	nan	nan	-12.047584	-172.014252	0	0	MERMAIDHydrophone(452.020)	nan	0.000000	nan
Measurement:depth:<absolute_pressure_gauge_name>	2018-08-06T14:54:20Z	MH	P0008	nan	nan	nan	nan	0	300	MERMAIDHydrophone(452.020)	nan	nan	nan
Measurement:depth:<absolute_pressure_gauge_name>	2018-08-06T15:54:20Z	MH	P0008	nan	nan	nan	nan	0	600	MERMAIDHydrophone(452.020)	nan	nan	nan
Measurement:depth:<absolute_pressure_gauge_name>	2018-08-06T16:54:20Z	MH	P0008	nan	nan	nan	nan	0	900	MERMAIDHydrophone(452.020)	nan	nan	nan
Measurement:depth:<absolute_pressure_gauge_name>	2018-08-06T17:54:20Z	MH	P0008	nan	nan	nan	nan	0	1200	MERMAIDHydrophone(452.020)	nan	nan	nan
Measurement:depth:<absolute_pressure_gauge_name>	2018-08-06T18:54:20Z	MH	P0008	nan	nan	nan	nan	0	1500	MERMAIDHydrophone(452.020)	nan	nan	nan
Algorithm:automaid:v3.5.0	2018-08-08T01:42:00Z	MH	P0008	0	BDH	-12.074427	-171.996506	0	1531	MERMAIDHydrophone(452.020)	20.0	nan	-0.290870
Measurement:depth:<absolute_pressure_gauge_name>	2018-08-14T20:42:00Z	MH	P0008	nan	nan	nan	nan	0	1500	MERMAIDHydrophone(452.020)	nan	nan	nan
Measurement:depth:<absolute_pressure_gauge_name>	2018-08-14T21:42:00Z	MH	P0008	nan	nan	nan	nan	0	1200	MERMAIDHydrophone(452.020)	nan	nan	nan
Measurement:depth:<absolute_pressure_gauge_name>	2018-08-14T22:42:00Z	MH	P0008	nan	nan	nan	nan	0	900	MERMAIDHydrophone(452.020)	nan	nan	nan
Measurement:depth:<absolute_pressure_gauge_name>	2018-08-14T23:42:00Z	MH	P0008	nan	nan	nan	nan	0	600	MERMAIDHydrophone(452.020)	nan	nan	nan
Measurement:depth:<absolute_pressure_gauge_name>	2018-08-15T00:00:00Z	MH	P0008	nan	nan	nan	nan	0	300	MERMAIDHydrophone(452.020)	nan	nan	nan
Measurement:GPS:u-blox_NEO-M8N	2018-08-15T00:10:52Z	MH	P0008	nan	nan	-12.205566	-171.903763	0	0	MERMAIDHydrophone(452.020)	nan	1.643707	nan
Measurement:GPS:u-blox_NEO-M8N	2018-08-15T00:12:58Z	MH	P0008	nan	nan	-12.205900	-171.904053	0	0	MERMAIDHydrophone(452.020)	nan	0.000000	nan
Measurement:GPS:u-blox_NEO-M8N	2018-08-15T00:23:25Z	MH	P0008	nan	nan	-12.207367	-171.905502	0	0	MERMAIDHydrophone(452.020)	nan	-0.000184	nan
Measurement:GPS:u-blox_NEO-M8N	2018-08-15T00:26:41Z	MH	P0008	nan	nan	-12.207784	-171.905991	0	0	MERMAIDHydrophone(452.020)	nan	0.000000	nan

Status of ESO effort in FDSN

- ESO is a “regular” member of the FDSN with Tim as representative and Frederik as alternate
- Framework proposal submitted to FDSN WGV to adopt GeoCSV format as an intermediary standard
- When approved, the GeoCSV info will be mapped into the intonational seismological metadata standard StationXML

Seismic data as a possible Essential Ocean Variable (EOV)

- EOVs are identified by GOOS Expert Panels based on:
- **Relevance:** The variable is effective in addressing the overall GOOS Themes – Climate, Operational Ocean Services, and Ocean Health.
- **Feasibility:** Observing or deriving the variable on a global scale is technically feasible using proven, scientifically understood methods.
- **Cost effectiveness:** Generating and archiving data on the variable is affordable, mainly relying on coordinated observing systems using proven technology, taking advantage where possible of historical datasets.
- Recommendation 1: ESO measurements should be made EOVs. With steering committee approval, the data committee will move this effort forward.



Comprehensive Test Ban Treaty Organization
- ESO should offer all of their data to CTBTO

https://www.ctbto.org/specials/vdec/

Recommendation 2: ESO should approach the CTBTO and make ESO data available to the CTBTO Virtual Data Exploitation Centre (VDEC)

Building the ESO Data System

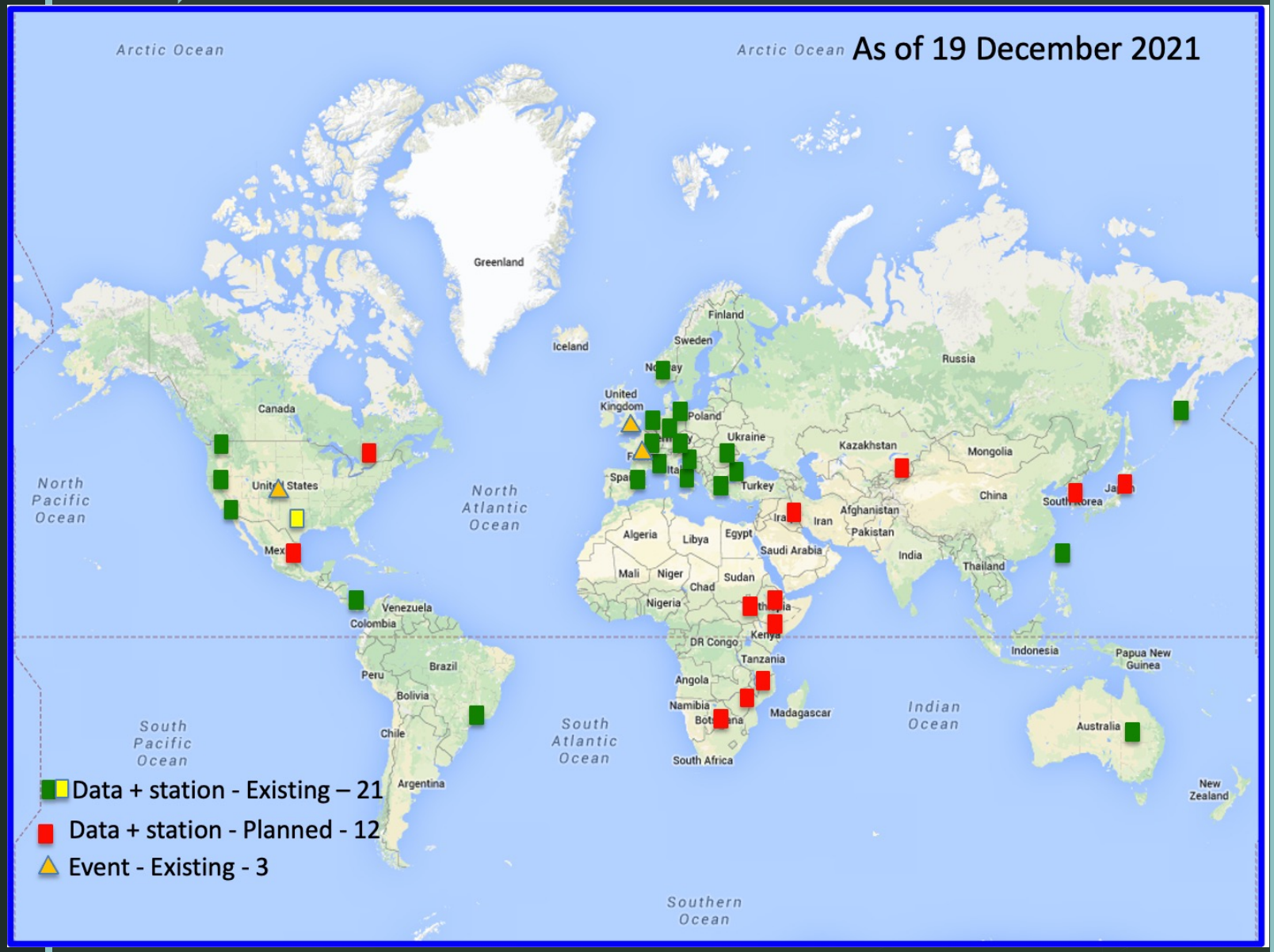
- Recommended Requirements for ESO data centres
- Recommended ESO Data Release policy

Recommended Requirements for ESO data centres

- ESO Data Centres must
  1. Support FDSN Standard data formats (SEED) and inbound transmission protocols (SeedLink) and mseed2dmc
  2. Support all FDSN Webservices for outbound data (e.g., Station, DataSelect, Event, and Availability)
  3. Participate in the FDSN federated data centres effort
  4. Routinely calculate data quality metrics following FDSN requirements (IRIS MUSTANG or EIDAWS-WFCatalog Service) for all MERMAID data they make publicly available.
  5. ESO Data Centres must provide rapidly varying metadata such as position in GeoCSV format

Recommendation 3: ESO supports the 5 requirements for an ESO Data Centre

Leverage the Existing Global Seismological Data Centre System



ESO Data Collection Centre

- ESO partners should consider funding a dedicated EarthScope Oceans Data Collection Centre ensuring proper format compliance before transmission to an ESO data centre.

Recommendation 4: ESO partners support the concept of a ESO Data Collection Centre

Recommend a Phased Approach

- Until these requirements are met, MERMAID data will go to Joel for processing and then be forwarded to the IRIS DMC which already supports all requirements
- An ESO component wishing to manage and distribute their own data could purchase the SeisComp3 system for use in their centres.
  - SeisComp3 is available from GEMPA in Potsdam, Germany (https://www.gempa.de).
  - For quality control, ESO Data Centres could use
    - EIDA Quality Tools (https://www.orfeus-eu.org/data/eida/quality/metrics/)
    - IRIS Mustang

Recommendation 5: ESO steering committee supports the concept of a phased approach for data centres that will not impede access to ESO data by the broader scientific community

Recommended ESO Data Release policy

- ESO Data Committee Recommends
  - ESO should support a single, consistent, ESO data release policy
- **Recommended Data Policy and Citation**
  - The data policy established by ESO is quite progressive. Data from at least 10% of an ESO partners MERMAIDS should be released without any delay other than the time required for data curation.
  - Data from all other stations must be released after two years.
  - Reference ESO data using DOI 10.7914/SN/MH.

For more details concerning early access to restricted data please refer to

https://geoweb.princeton.edu/people/simons/earthscopeoceans/data/metadata.html.

Recommendation 6: ESO steering committee supports the recommended ESO Data Release Policy

Requested Steering Committee Approvals

1. Recommendation 1: ESO measurements should be made EOVs. With steering committee approval, the data committee will move this effort forward.
2. Recommendation 2: ESO should approach the CTBTO and make ESO data available to the CTBTO Virtual Data Exploitation Centre (VDEC)
3. Recommendation 3: ESO supports the 5 requirements for an ESO Data Centre
4. Recommendation 4: ESO partners support the concept of a ESO Data Collection Centre
5. Recommendation 5: ESO steering committee supports the concept of a phased approach for data centres that will not impede access to ESO data by the broader seismological community
6. Recommendation 6: ESO supports the recommended ESO Data Release Policy
7. Recommendation: Add another Data Committee member from SUSTECH
    1. With John Chen's concurrence and recommendation