

Vertical CRS Storage in BAG Metadata

ONSWG Position Paper 2022-01

2022-05-19

# Problem Statement

The Bathymetric Attributed Grid (BAG) file format has, in the embedded XML-formatted ISO 19115 metadata, specification for coordinate reference systems (CRS), defined through OGC Well Known Text (WKT) strings as appropriate for the PROJ coordinate transformation package (and consequently for libraries such as the Geographic Data Abstraction Library (GDAL), which are commonly used in geospatial software). A separate definition is provided for the horizontal CRS, and for the vertical. Both use gmd:referenceSystemInfo metadata element encodings (gmd:referenceSystemInfo > gmd:MD\_ReferenceSystem > gmd:referenceSystemIdentifier > gmd:RS\_Identifier > gmd:code > gco:CharacterString).

The horizontal CRS specification is generally well understood and supported, but the vertical CRS has received less attention in the past, primarily because most data was specified on chart datum, which the complexity of this dealt with through charting practice, rather than being specified completely in the file.

With modern surveys being conducted more frequently on a constructed or mathematical CRS, most commonly the relevant ellipsoid, more precise definition of the vertical CRS, how the data came to be converted onto that CRS, and how it may be related to other plausible CRS need to be encoded into the product of the survey, in our case the BAG file. Among other issues, outstanding technical questions include:

1. Precise definition of an “understood” datum. For example, a Mean Lower Low Water (MLLW) vertical CRS would require some definition of the agency making the definition, the tidal epoch to which it refers, and the location.
2. Archive of the steps taken to convert to a given datum. For example, if the data were collected to the ellipsoid, but then converted to MLLW or Lowest Astronomical Tide (LAT) for final product, knowing how the conversion was done would be essential to undoing the transformation prior to applying another.
3. For some transformation methods (for example NOAA’s VDATUM), the transformation is defined in terms of a number of intermediate datums. An appropriate record of how the transformation was constructed is essential metadata that needs to be preserved.
4. An axiomatic definition for BAG files is that the vertical is positive up (i.e., elevation rather than conventional hydrographic practice of depth below the water surface being positive). A unambiguous statement of this convention is required to ensure that third-party libraries (e.g., Proj.4 or GDAL) correctly interpret the data while reading.

There are a number of different solutions to this, including WKT 1.0, WKT 2.0, and EPSG codes. While WKT strings are readily constructed, and need no authorization, they are essentially free text, and therefore can be interpreted differently by different users and data constructors. EPSG codes are better defined since they come from a controlled registry, but registering new codes can take some time. Current field practice appears to have a mixture of WKT 2.0 text with additional free text “comments” with extra information (which are difficult to interpret cleanly), EPSG codes, and WKT 1.0 text without extra definitions. This mixture means that libraries to read BAG files have to be more complex, and/or make guesses as to interpretation. This is directly contrary to the principle of axiomatic definition inherent in BAG files, specifically to reduce this type of complexity.

The Working Group therefore needs to determine a single technical solution to the problem of vertical CRS specification and make recommendations for its adoption. In order to preserve backwards compatibility, the WG should also recommend a version change for the BAG specification.

# Potential Solutions for the Specification of the Vertical CRS

## Option 1: Controlled Vocabulary

As a simple solution, the WG could specify a number of known vertical CRS that were tied to known standards and provide simply a text description of which CRS is being used. Although simple, this would very likely result in a significant maintenance burden, and potentially the requirement for a very large number of “known” vertical CRS.

## Option 2: WKT 2.0

Unlike WKT 1.0, where there is definition, but significant latitude in the text, WKT 2.0 is more closely defined, and has many more options for specific definition of complex vertical CRS. Selection of WKT 2.0 (https://www.ogc.org/standards/wkt-crs) as an encoding, along with the correct specification of this choice (gmd:codeSpace), should allow for more detailed description of the CRS.

## Option 3: EPSG Codes

EPSG codes (https://epsg.io) have the benefit of being defined precisely in a standard repository (e.g., Option 1), and would be maintained over a wide range of options without effort from the WG, but would also require effort to register if a CRS is used which is not available in the standard registry (which might take some time).

## Option 4: Auxiliary Optional Separation Layer

Due to work by Leidos, BAG files have an optional layer for separation grids, which could be used to memorialize the (possibly composite) offset between a known, well defined, CRS and the CRS in which the BAG data is represented. For example, it could provide an offset from the ITRF2000-derived ellipsoid to an arbitrary tidal datum in a manner similar to SHOM’s Bathyelli grids. This would require only a single vertical CRS definition to be maintained (the standard reference datum), with the responsibility being with the data creator to provide the composite offset grid. It would, however, also mean more data being recorded in every BAG file (a disadvantage), and would only provide a one-step offset, without recording the various components of the transformation (e.g., as with NOAA VDATUM).

## Auxiliary Requirement: Lineage of Transformation

Maintaining metadata on transformations of vertical CRS may be required. The BAG metadata has a lineage segment (gmd:dataQualityInfo > gmd:DQ\_DataQuality > gmd:lineage), although this has been rarely used. Some guidance on encoding of information in the lineage section for transformation details (or CRS details) would be required if this mechanism were to be used.

# Decisions Required

The WG should conclude on:

1. A single technical solution for the specification of vertical CRS metadata.
2. A timeline for the adoption of the selected technical solution.
3. A responsible party (or parties) for any library maintenance to support the technical solution selected.
4. A new version number for the BAG file format to allow the library to know what to expect from the vertical CRS metadata entry.

# References

BAG Metadata: <https://github.com/OpenNavigationSurface/BAG/wiki/FSD-BAGStructure#metadata>

WKT1 Specification: <https://docs.opengeospatial.org/is/12-063r5/12-063r5.html>

WKT2 Specification: <https://docs.opengeospatial.org/is/18-010r7/18-010r7.html>

EPSG: https://epsg.org/

PROJ: https://proj.org/

GDAL: <https://gdal.org/>

SHOM Bathyelli Grids : (http://refmar.shom.fr/applications\_maregraphiques/programmes-projets/projets/bathyelli).