
Appendix: WaterML – WaterML XML Time Series File Input Type

2015-09-14

Overview

The WaterML input type corresponds to the WaterML time series file format. WaterML is now an Open Geospatial Consortium (OGC) standard. See the following resources:

- OGC WaterML website: <http://www.opengeospatial.org/standards/waterml>
- Kisters WaterML 2.0 website: <http://www.waterml2.org/>
- CUAHSI WaterOneFlow web services website: <http://his.cuahsi.org/wofws.html>

WaterML 1.1 and 2.0 formats are supported at various levels in the following TSTool commands: `ReadUsgsNwisDaily()`, `ReadUsgsNwisDaily()`, `ReadUsgsNwisInstantaneous()`, `ReadWaterML()`, and `WriteWaterML()`. The `WebGet()` command also can be used to retrieve WaterML files from web services such as the USGS NWIS site (see the **UsgsNwisDaily Data Store** appendix).

WaterML and Standard Time Series Properties

This section needs to be updated to discuss WaterML 2.0. The standard time series identifier for WaterML time series in TSTool is of the form:

`Location.DataSource.DataType.Interval~InputType~InputName`

More specifically, the identifier adheres to the following convention:

`SiteCode.AgencyCode.VariableCode-StatisticCode.Interval~WaterML~InputFile`

where identifier parts are described as follows:

- The `SiteCode` corresponds to the `siteCode` XML element in the `sourceInfo` element (e.g., this is the USGS stream gage station identifier).
- `AgencyCode` corresponds to the `agencyCode` attribute of the `siteCode` XML element (e.g., USGS for USGS NWIS stream gage stations).
- `VariableCode` corresponds to the `variableCode` XML element in the `variable` element (e.g., this is 00060 for streamflow for USGS stream gages). In the future, a more humanly-readable text value may be used, perhaps with a string prefix; however, such handling must guarantee that the variable is unique.
- `StatisticCode` corresponds to the `optionCode` attribute of the `option` XML element in the `options` element, in the `variable` element (e.g., 00006 for USGS NWIS sum statistic). In the future, a more humanly-readable text value may be used, perhaps with a string prefix; however, such handling must guarantee that the variable is unique.
- According to the WaterML specification, `Interval` should correspond to the `timeScale` or `timeSpacing` XML elements. However, current USGS NWIS WaterML does not include this information and the interval must be specified externally when reading the WaterML. A more

thorough handling of the interval will be explored as support for WaterOneFlow WaterML web services is added.

- InputType is WaterML and InputName is the name of the WaterML file.
- Data units correspond to the unitCode XML element in the unit element of the variable element
- Missing numerical values are internally represented as NaN and are assigned to any date/times in the period that do not have values. The noDataValue element in the variable element is checked and matching data values are handled as missing. Often, however, missing values are simply not included as data.
- Data value flags, if encountered, are retained in the time series. However, because the USGS uses “A” for approved, it may be necessary to ignore this flag so that other flags stand out more when visualized.
- Data value qualifiers definitions are saved with time series and are available to use in time series visualization as flag definitions

Limitations

WaterML limitations relative to TSTool standard features are as follows:

- WaterML files from NWIS do not indicate the interval of the data. Day can be assumed for the daily values web service; however, trying to read the WaterML file later requires that the interval is specified.
- Some of the descriptions contain units, which may lead to confusion if time series are processed into different units.

TSTool Configuration File

WaterML support in TSTool is enabled/disabled using a configuration property as shown below.

```
# Configuration file for TSTool

[TSTool]

WaterMLEnabled = true
```

TSTool Configuration File with WaterML Properties